Publ. Mat. **51** (2007), 397–456

WEIGHTED NORM INEQUALITIES FOR CALDERÓN-ZYGMUND OPERATORS WITHOUT DOUBLING CONDITIONS

Xavier Tolsa

Abstract

Let μ be a Borel measure on $\mathbb{R}^d$ which may be non doubling. The only condition that μ must satisfy is $\mu(B(x,r)) \leq Cr^n$ for all $x \in \mathbb{R}^d$, $r > 0$ and for some fixed n with $0 < n \leq d$. In this paper we introduce a maximal operator N, which coincides with the maximal Hardy-Littlewood operator if $\mu(B(x,r)) \approx r^n$ for $x \in \operatorname{supp}(\mu)$, and we show that all n-dimensional Calderón-Zygmund operators are bounded on $L^p(w\,d\mu)$ if and only if N is bounded on $L^p(w\,d\mu)$, for a fixed $p \in (1,\infty)$. Also, we prove that this happens if and only if some conditions of Sawyer type hold. We obtain analogous results about the weak (p,p) estimates. This type of weights do not satisfy a reverse Hölder inequality, in general, but some kind of self improving property still holds. On the other hand, if $f \in RBMO(\mu)$ and $\varepsilon > 0$ is small enough, then $e^{\varepsilon f}$ belongs to this class of weights.

1. Introduction

Let μ be some Borel measure on $\mathbb{R}^d$ satisfying

$$\mu(B(x,r)) \leq C_0 r^n \quad \text{for all } x \in \mathbb{R}^d,\ r > 0, \tag{1.1}$$

where n is some fixed constant (which may be non integer) with $0 < n \leq d$. In this paper we obtain a characterization of all the weights w such that, for every n-dimensional Calderón-Zygmund operator (CZO) T which is bounded on $L^2(\mu)$, the following weighted inequality holds:

$$\int |Tf|^p\, w\, d\mu \leq C \int |f|^p\, w\, d\mu \quad \text{for all } f \in L^p(w), \tag{1.2}$$

2000 *Mathematics Subject Classification.* Primary: 42B20; Secondary: 42B25.
Key words. Calderón-Zygmund operators, weights, non doubling measures.
Partially supported by grants MTM2004-00519 and Acción Integrada HF2004-0208 (Spain), and 2001-SGR-00431 (Generalitat de Catalunya).

where C is independent of f, $1 < p < \infty$, and $L^p(w) := L^p(w\, d\mu)$. It is shown that these weights w are those such that a suitable maximal operator N (defined below) is bounded on $L^p(w)$. We also prove an analogous result for the weak (p, p) estimates.

Moreover, we show that the L^p weights for CZO's (and for N) satisfy a self improving property. Loosely speaking, weak weighted inequalities for w and for the dual weight $w^{-1/(p-1)}$ imply strong weighted inequalities for w and its dual weight. Let us remark that we do *not* assume that the underlying measure μ is doubling. Recall that μ is said to be doubling if there exists some constant C such that $\mu(B(x, 2r)) \leq C\, \mu(B(x, r))$ for all $x \in \operatorname{supp}(\mu)$ and $r > 0$.

In the particular case where μ coincides with the Lebesgue measure on $\mathbb{R}^d$, it is known that the weighted inequality (1.2) holds for all d-dimensional CZO's if and only if w is an A_p weight. This result was obtained by Coifman and Fefferman [**CF**], and it generalizes a previous result by Hunt, Muckenhoupt and Wheeden [**HMW**] about the Hilbert transform. Let us recall that Muckenhoupt proved [**Mu**] that the A_p weights are precisely those weights w for which the Hardy-Littlewood operator is bounded on $L^p(w)$ (always assuming μ to be the Lebesgue measure on $\mathbb{R}^d$). So the L^p weights for CZO's and the L^p weights for the maximal Hardy-Littlewood operator coincide in this case.

Suppose now that the measure μ satisfies

$$\mu(B(x, r)) \approx r^n \quad \text{for all } x \in \operatorname{supp}(\mu),\ r > 0, \tag{1.3}$$

where $A \approx B$ means that there is some constant $C > 0$ such that $C^{-1}A \leq B \leq CA$, with C depending only on n and d (and also on C_0 sometimes), in general. In this case the results (and their proofs) are analogous to the ones for the Lebesgue measure. Namely, (1.2) holds for all n-dimensional CZO's if and only if $w \in A_p$, which is equivalent to say that the maximal Hardy-Littlewood operator is bounded on $L^p(w)$.

Many other results about weights for CZO's can be found in the literature. In most of them it is assumed that μ is either the Lebesgue measure on $\mathbb{R}^d$ or the underlying measure of a space of homogeneous type, satisfying (1.3). See for example [**Pé**] and the recent result on the two weight problem for the Hilbert transform in [**Vo**].

It is much more difficult to find results where (1.3) does not hold. Saksman [**Sak**] has obtained some results concerning the weights for the Hilbert transform H on arbitrary bounded subsets of $\mathbb{R}$ (with μ being the Lebesgue measure restricted to these subsets). These results relate the boundedness of H on $L^p(w)$ with some operator properties of H, and quite often his arguments are of complex analytic nature.

Orobitg and Pérez [**OP**] have studied the A_p classes of weights with respect to arbitrary measures on $\mathbb{R}^d$, which may be non doubling. In particular, they have shown that if w is an A_p weight, then the centered maximal Hardy-Littlewood operator is bounded on $L^p(w)$, and that if moreover μ satisfies (1.1), then all n-dimensional CZO's are also bounded on $L^p(w)$. Other more recent result which involve the operator

$$M_k f(x) = \sup_{x\in B} \frac{1}{\mu(kB)} \int_B |f(y)|\, d\mu(y) \quad \text{for } x \in \operatorname{supp}(\mu) \text{ and } k > 1,$$

where the supremum is taken over all balls B containing x, have been obtained in [**Ko**].

Our approach uses real variable techniques and it is based on the ideas and methods developed in [**To2**], [**To3**] and [**To4**] to extend Calderón-Zygmund theory to the the setting of non doubling measures. Indeed, recently it has been shown that the doubling assumption is not essential for many results of Calderón-Zygmund theory. See [**NTV1**], [**To1**], [**NTV2**], [**MMNO**] and [**GM**], for instance, in addition to the references cited above.

In order to state our results more precisely, we need to introduce some definitions. A kernel $k(\cdot,\cdot)\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is called a (n-dimensional) Calderón-Zygmund (CZ) kernel if

(1) $|k(x,y)| \leq \dfrac{C_1}{|x-y|^n}$ if $x \neq y$,

(2) there exists some fixed constant $0 < \gamma \leq 1$ such that

$$|k(x,y) - k(x',y)| + |k(y,x) - k(y,x')| \leq C_2 \frac{|x-x'|^\gamma}{|x-y|^{n+\gamma}}$$

if $|x-x'| \leq |x-y|/2$.

Throughout all the paper we will assume that μ is a Radon measure on $\mathbb{R}^d$ satisfying (1.1). We say that T is a (n-dimensional) CZO associated to the kernel $k(x,y)$ if for any compactly supported function $f \in L^2(\mu)$

$$Tf(x) = \int k(x,y)\, f(y)\, d\mu(y) \quad \text{if } x \notin \operatorname{supp}(\mu), \tag{1.4}$$

and T is bounded on $L^2(\mu)$ (see the paragraph below regarding this question). If we want to make explicit the constant γ which appears in the second property of the CZ kernel, we will write $T \in CZO(\gamma)$.

The integral in (1.4) may be non convergent for $x \in \operatorname{supp}(\mu)$, even for "very nice" functions, such as $\mathcal{C}^\infty$ functions with compact support.

For this reason it is convenient to introduce the truncated operators T_ε, $\varepsilon > 0$:

$$T_\varepsilon f(x) = \int_{|x-y|>\varepsilon} k(x,y)\, f(y)\, d\mu(y).$$

Then we say that T is bounded on $L^2(\mu)$ if the operators T_ε are bounded on $L^2(\mu)$ uniformly on $\varepsilon > 0$.

Now we will define the maximal operator N. For $0 < r < R$ and a fixed $x \in \operatorname{supp}(\mu)$, we consider the function

$$\varphi_{x,r,R}(y) = \begin{cases} 1/r^n & \text{if } 0 \le |x-y| \le r, \\ 1/|x-y|^n & \text{if } r \le |x-y| \le R, \\ 0 & \text{if } |x-y| > R. \end{cases}$$

Then we set

$$Nf(x) = \sup_{0<r<R} \frac{1}{1 + \|\varphi_{x,r,R}\|_{L^1(\mu)}} \int |\varphi_{x,r,R}\, f|\, d\mu, \tag{1.5}$$

for $f \in L^1_{\mathrm{loc}}(\mu)$ and $x \in \operatorname{supp}(\mu)$.

Throughout all the paper w stands for a positive weight in $L^1_{\mathrm{loc}}(\mu)$. Sometimes the measure $w\, d\mu$ is denoted simply by w. The notation for the dual weight is $\sigma := w^{-1/(p-1)}$, with $1 < p < \infty$.

The first result that we will prove deals with the weak (p,p) estimates.

Theorem 1.1. *Let p, γ be constants with $1 \le p < \infty$ and $0 < \gamma \le 1$. Let w be a positive weight. The following statements are equivalent:*

(a) *All operators $T \in CZO(\gamma)$ are of weak type (p,p) with respect to $w\, d\mu$.*

(b) *The maximal operator N is of weak type (p,p) with respect to $w\, d\mu$.*

Next we state the corresponding result for the strong (p,p) estimates.

Theorem 1.2. *Let p, γ be constants with $1 < p < \infty$ and $0 < \gamma \le 1$. Let w be a positive weight. The following statements are equivalent:*

(a) *All operators $T \in CZO(\gamma)$ are bounded on $L^p(w)$.*

(b) *The maximal operator N is bounded on $L^p(w)$.*

Let us denote by Z_p the class of weights w such that N is bounded on $L^p(w)$, and by Z_p^{weak} its weak version, that is, the class of weights w such that N is bounded from $L^p(w)$ into $L^{p,\infty}(w)$. Notice that since N is bounded on $L^\infty(w)$, by interpolation we have $Z_p \subset Z_q$ if $1 < p \le q < \infty$. On the other hand, the inclusion $Z_p \subset Z_p^{\mathrm{weak}}$ is trivial, and by duality (of CZO's) and Theorem 1.2 it follows that $w \in Z_p$ if and only if $\sigma \in Z_{p'}$, where p' stands for the conjugate exponent of p, i.e. $p' = p/(p-1)$.

We will prove the following self improving property for this type weights:

Theorem 1.3. *Let w be a positive weight and $1 < p < \infty$. If $w \in Z_p^{\text{weak}}$ and $\sigma = w^{-1/(p-1)} \in Z_{p'}^{\text{weak}}$, then $w \in Z_p$ and $\sigma \in Z_{p'}$.*

More detailed results are stated in Lemmas 4.1 and 4.2 in Section 4. In particular, necessary and sufficient conditions of "Sawyer type" are given for the boundedness of N on $L^p(w)$ and also for the weak (p,p) case. Moreover, it is shown that if $w \in Z_p$ ($w \in Z_p^{\text{weak}}$), then the maximal CZO

$$T_* f(x) = \sup_{\varepsilon>0} |T_\varepsilon f(x)|$$

is bounded on $L^p(w)$ [of weak type (p,p) with respect to w].

Let us see an easy consequence of our results. Given $\lambda \geq 1$, let M_λ be following version of the maximal Hardy-Littlewood operator:

$$(1.6)\quad M_\lambda f(x) = \sup_{r>0} \frac{1}{\mu(B(x,\lambda r))} \int_{B(x,r)} |f|\, d\mu, \quad f \in L^1_{\text{loc}}(\mu),\ x \in \operatorname{supp}(\mu).$$

It is easily seen that for any $\lambda \geq 1$,

$$(1.7)\qquad Nf(x) \leq C(\lambda)\, M_\lambda f(x), \quad f \in L^1_{\text{loc}}(\mu),\ x \in \operatorname{supp}(\mu).$$

Thus all weights w such that M_λ is bounded on $L^p(w)$ belong to Z_p, and then all CZO's are bounded on $L^p(w)$. In particular, $A_p \subset Z_p$ if $1 < p < \infty$.

Observe that the maximal operator N is a centered maximal operator, which is not equivalent to any "reasonable" non centered maximal operator, as far as we know. This fact and the absence of any doubling condition on μ are responsible for most of the difficulties that arise in our arguments. For instance, it turns out that the weights of the class Z_p don't satisfy a reverse Hölder inequality, in general. Indeed there are examples which show that it may happen that $w \in Z_p$ but $w^{1+\varepsilon} \notin L^1_{\text{loc}}(\mu)$ for any $\varepsilon > 0$ (see Examples 2.3 and 2.4). Also, we will show that the weights in Z_p satisfy a property much weaker than the A_∞ condition of the classical A_p weights (see Definition 5.1 and Lemma 5.3), which is more difficult to deal with.

Let us notice that it has been shown in [**OP**] that, even with μ non doubling, if $w \in A_p$, then w satisfies a reverse Hölder inequality. As a consequence, $A_p \neq Z_p$ in general.

The plan of the paper is the following. In Section 2 we show some examples which illustrate our results. In Section 3 we recall the basic properties of the lattice of cubes introduced in [**To3**] and [**To4**], together with its associated approximation of the identity. This construction will be an essential tool for our arguments. In the same section we will study some of the properties of the maximal operator N. In Section 4 we state Lemmas 4.1 and 4.2, from which Theorems 1.1, 1.2 and 1.3 follow directly. Lemma 4.1 deals with the weak (p,p) estimates, and it is proved in Sections 5–7, while the strong (p,p) case is treated in Lemma 4.2 and is proved in Sections 8–10. In Section 11 we explain how to prove the theorems above in their full generality, without a technical assumption that is used in Sections 5–10 for simplicity. Finally, in Section 12 we show which is the relationship between Z_p and $RBMO(\mu)$ (this is the space of type BMO introduced in [**To2**]), and we make some remarks. In particular in this section we prove the following result:

Theorem 1.4. *Let* $1 < p < \infty$. *If* $f \in RBMO(\mu)$ *and* $\varepsilon = \varepsilon(\|f\|_*, p) > 0$ *is small enough, then* $e^{\varepsilon f} \in Z_p$.

For the precise definition of $RBMO(\mu)$, see Section 12.

2. Some examples

Example 2.1. If $\mu(B(x,r)) \approx r^n$ for all $x \in \operatorname{supp}(\mu)$, then $Nf(x) \approx Mf(x)$, where M is the usual centered Hardy-Littlewood operator (defined in (1.6) with $\lambda = 1$). In this case, the class Z_p coincides with the class A_p.

Example 2.2. In $\mathbb{R}^2$, consider the square $Q_0 = [0,1]^2$ and the measure $d\mu = \chi_{Q_0}\, dm$, where dm stands for the planar Lebesgue measure, and take $n = 1$. That is, we are interested in studying the weights for 1-dimensional CZO's such as the Cauchy transform. Notice that μ is a doubling measure which does not satisfy the assumption in Example 2.1. For this measure, we have the uniform estimate $\int \frac{1}{|y-x|}\, d\mu(y) \leq C$. Then, from Theorem 1.2, we deduce that the class Z_p coincides with the class of L^p weights for the fractional integral

$$I_0 f(x) = \int \frac{1}{|y-x|} f(y)\, d\mu(y),$$

since $Nf(x) \approx I_0|f|(x)$. This is the result that should be expected because, with our choice of μ, I_0 is a CZO, and for all other $T \in CZO(\gamma)$, we have $|Tf(x)| \leq C_1\, I_0|f|(x)$.

Example 2.3. This is an example studied by Saksman in his paper about weights for the Hilbert transform [**Sak**]. We are in $\mathbb{R}$ and $n = 1$. Let $\ell_k = 1/k!$ and consider the intervals $I_k = \left(\frac{1}{k} - \frac{\ell_k}{4}, \frac{1}{k} + \frac{\ell_k}{4}\right)$ for $k \geq 1$. Let μ be the Lebesgue measure restricted to the set $S := \bigcup_{k=1}^{\infty} I_k$.

Let w be a weight such that $w \geq 1$ and $w_{|I_k}$ is constant for each $k \geq 1$. In [**Sak**], it is proved that, for any $p \in (1, \infty)$, the Hilbert transform is bounded on $L^p(w)$ if and only if $w \in L^1(\mu)$. Almost the same calculations show that the operators S_k (defined after Lemma 3.7 below) are uniformly bounded on $L^p(w)$ if and only if $w \in L^1(\mu)$.

So, if a weight w_0 is defined by $w_{0|I_k} = (n-2)!$, then $w_0 \in Z_p$ for all $p \in (1, \infty)$, by Lemma 4.2 below. However, it is easily seen that $w_0^{1+\varepsilon} \notin L^1(\mu)$ for any $\varepsilon > 0$. Therefore, w_0 does not satisfy a reverse Hölder inequality.

Example 2.4. In this example we will show that there are measures μ and weights $w \in Z_p$ such that the (centered) maximal Hardy-Littlewood operator M is not bounded on $L^p(w)$. Also we will see that it may happen $w \in Z_p$ but $w \notin Z_{p-\varepsilon}$ for any $\varepsilon > 0$.

We take $d = n = 1$. Suppose that I_1 and I_2 are disjoint intervals on $\mathbb{R}$. The measure μ is the Lebesgue measure restricted to $I_1 \cup I_2$. Suppose that $\mu(I_1) = \mu(I_2) = L$, and let $D = \operatorname{dist}(I_1, I_2)$, with $D \geq 2L$. For $f = \chi_{I_1}$, the inequality $\|Mf\|_{L^p(w)} \leq C_3 \|f\|_{L^p(w)}$ implies

$$w(I_2) \leq C_4\, w(I_1), \tag{2.1}$$

with C_4 depending on C_3 but not on D or L. By symmetry, (2.1) also holds interchanging I_1 and I_2.

Also, if $w \in Z_p$, from $\|Nf\|_{L^p(w)} \leq C_5 \|f\|_{L^p(w)}$ we get $\frac{L}{D} w(I_2)^{1/p} \leq Cw(I_1)^{1/p}$. That is,

$$\left(\frac{L}{D}\right)^p w(I_2) \leq C_6\, w(I_1). \tag{2.2}$$

The constant C_6 depends only on C_5. By symmetry, we deduce

$$C_6^{-1} \left(\frac{L}{D}\right)^p w(I_2) \leq w(I_1) \leq C_6 \left(\frac{L}{D}\right)^{-p} w(I_2). \tag{2.3}$$

If w is constant on each interval I_1, I_2, then N is bounded on $L^p(w)$ and it is easily seen that $\|N\|_{L^p(w) \to L^p(w)} \leq C(C_6)$.

Now we introduce a new measure μ on $\mathbb{R}$. For each integer $m \geq 1$ we consider the intervals $I_1^m = 1000^m + [-m-1, -m]$ and $I_2^m = 1000^m + [m, m+1]$, so that $D_m := \operatorname{dist}(I_1^m, I_2^m) = 2m$ and $L = 1$. The measure μ is the Lebesgue measure restricted to $\bigcup_{m=1}^\infty (I_1^m \cup I_2^m)$. The weight w is constant on each interval $w_{I_j^m}$, with $w_{|I_1^m} \equiv 1$ and $w_{|I_2^m} \equiv (D_m/L)^p = (2m)^p$.

The maximal operator M is not bounded on $L^p(w)$ because, otherwise, we should have $w(I_2^m) \leq C\, w(I_1^m)$ uniformly on m, as in (2.1). On the other hand, (2.3) is satisfied (with the corresponding subindices and superindices m) uniformly on m. Taking also into account that $I_1^m \cup I_2^m$ is very far from $I_1^r \cup I_2^r$ if $m \neq r$, it is easily checked that N is bounded on $L^p(w)$. Moreover, N is not bounded on $L^{p-\varepsilon}(w)$ for any $\varepsilon > 0$ because the inequality

$$\left(\frac{L}{D}\right)^{p-\varepsilon} w(I_2) \leq C\, w(I_1)$$

fails for m big enough.

3. Preliminaries

3.1. The lattice of cubes. For definiteness, by a cube we mean a closed cube with sides parallel to the coordinate axes. We will assume that the constant C_0 in (1.1) has been chosen big enough so that for all cubes $Q \subset \mathbb{R}^d$ we have $\mu(Q) \leq C_0\, \ell(Q)^n$, where $\ell(Q)$ stands for the side length of Q.

Given $\alpha, \beta > 1$, we say that the cube $Q \subset \mathbb{R}^d$ is (α,β)-doubling if $\mu(\alpha Q) \leq \beta\mu(Q)$. If α and β are not specified and we say that some cube is doubling, we are assuming $\alpha = 2$ and β equal to some constant big enough ($\beta > 2^d$, for example) which may depend from the context.

Remark 3.1. Due to the fact that μ satisfies the growth condition (1.1), there are a lot "big" doubling cubes. To be precise, given any point $x \in \operatorname{supp}(\mu)$ and $c > 0$, there exists some (α,β)-doubling cube Q centered at x with $l(Q) \geq c$. This follows easily from (1.1) and the fact that we are assuming that $\beta > \alpha^n$.

On the other hand, if $\beta > \alpha^d$, then for μ-a.e. $x \in \mathbb{R}^d$ there exists a sequence of (α,β)-doubling cubes $\{Q_k\}_k$ centered at x with $\ell(Q_k) \to 0$ as $k \to \infty$. So there are a lot of "small" doubling cubes too.

Given cubes Q, R, with $Q \subset R$, we denote by z_Q the center of Q, and by Q_R the smallest cube concentric with Q containing Q and R. We set

$$\delta(Q,R) = \int_{Q_R \setminus Q} \frac{1}{|x - z_Q|^n}\, d\mu(x).$$

We may treat points $x \in \mathbb{R}^d$ and the whole space $\mathbb{R}^d$ as if they were cubes (with $\ell(x) = 0$, $\ell(\mathbb{R}^d) = \infty$). So for $x \in \mathbb{R}^d$ and some cube Q, the notations $\delta(x,Q)$, $\delta(Q,\mathbb{R}^d)$ make sense. Of course, it may happen $\delta(x,Q) = \infty$ and $\delta(Q,\mathbb{R}^d) = \infty$.

In the following lemma, proved in [**To3**], we recall some useful properties of $\delta(\cdot,\cdot)$.

Lemma 3.2. *Let $P, Q, R \subset \mathbb{R}^d$ be cubes with $P \subset Q \subset R$. The following properties hold:*

(a) *If $\ell(Q) \approx \ell(R)$, then $\delta(Q, R) \leq C$. In particular, $\delta(Q, \rho Q) \leq C_0\, 2^n \rho^n$ for $\rho > 1$.*

(b) *If $Q \subset R$ are concentric and there are no doubling cubes of the form $2^k Q$, $k \geq 0$, with $Q \subset 2^k Q \subset R$, then $\delta(Q, R) \leq C_7$.*

(c) $\delta(Q, R) \leq C\left(1 + \log \dfrac{\ell(R)}{\ell(Q)}\right)$.

(d) $\big|\delta(P, R) - [\delta(P, Q) + \delta(Q, R)]\big| \leq \varepsilon_0$. *That is, with a different notation, $\delta(P, R) = \delta(P, Q) + \delta(Q, R) \pm \varepsilon_0$.*

The constants C and ε_0 that appear in (b), (c) and (d) depend on C_0, n, d. The constant C in (a) depends, further, on the constants that are implicit in the relation $\approx$. Let us insist on the fact that a notation such as $a = b \pm \varepsilon$ does not mean any precise equality but the estimate $|a - b| \leq \varepsilon$.

Now we will describe the lattice of cubes introduced in [**To4**]. In the following lemma, $Q_{x,k}$ stands for a cube centered at x, and we allow $Q_{x,k} = x$ and $Q_{x,k} = \mathbb{R}^d$. If $Q_{x,k} \neq x, \mathbb{R}^d$, we say that $Q_{x,k}$ is a *transit cube.*

Lemma 3.3. *Let A be an arbitrary positive constant big enough. There exists a family of cubes $Q_{x,k}$, for all $x \in \operatorname{supp}(\mu)$, $k \in \mathbb{Z}$, centered at x, and such that:*

(a) *$Q_{x,k} \subset Q_{x,j}$ if $k \geq j$.*

(b) *$\lim_{k\to+\infty} \ell(Q_{x,k}) = 0$ and $\lim_{k\to-\infty} \ell(Q_{x,k}) = \infty$.*

(c) *$\delta(Q_{x,k}, Q_{x,j}) = (j-k)A \pm \varepsilon$ if $j > k$ and $Q_{x,k}$, $Q_{x,j}$ are transit cubes.*

(d) *$\delta(Q_{x,k}, Q_{x,j}) \leq (j - k)A + \varepsilon$ if $j > k$.*

(e) *If $2Q_{x,k} \cap 2Q_{y,k} \neq \varnothing$, then $2Q_{x,k} \subset Q_{y,k-1}$ and $\ell(Q_{x,k}) \leq \ell(Q_{y,k-1})/100$.*

(f) *There exists some $\eta > 0$ such that if $m \geq 1$ and $2Q_{x,k+m} \cap 2Q_{y,k} \neq \varnothing$, then $\ell(Q_{x,k+m}) \leq 2^{-\eta A m} \ell(Q_{y,k})$.*

The constants ε, η in (c), (d) and (f) depend on C_0, n, d, but not on A.

See [**To4**, Section 3] for the proof. The constant ε above must be understood as an error term, because we will take $A \gg \varepsilon$. Let us notice also that, if necessary, the cubes $Q_{x,k}$ can be chosen so that they are doubling (see [**To4**]). However we don't need this assumption.

Remark 3.4. If $x \in \operatorname{supp}(\mu)$ is such that $\int_{B(x,1)} |y-x|^{-n}\, d\mu(y) < \infty$, then it follows from the properties of the lattice that there exists some $K_x \in \mathbb{Z}$ such that $Q_{x,k} = x$ for $k > K_x$ and $Q_{x,k} \neq x$ for $k \leq K_x$. In this case we say that $Q_{x,k}$ is a *stopping cube* (or *stopping point*).

If $\int_{\mathbb{R}^d \setminus B(x,1)} |y-x|^{-n}\, d\mu(y) < \infty$ (which does not depend on $x \in \operatorname{supp}(\mu)$), then there exists some constant $\bar{K}_x$ such that $Q_{x,k} = \mathbb{R}^d$ for $k < \bar{K}_x$ and $Q_{x,k} \neq \mathbb{R}^d$ for $k \geq \bar{K}_x$. We say that $\mathbb{R}^d$ is an (or the) *initial cube*. From the property (e) in the lemma above, it follows easily that $|\bar{K}_x - \bar{K}_y| \leq 1$ for $x, y \in \operatorname{supp}(\mu)$. However, as shown in [**To4**], the construction of the lattice can be done so that $\bar{K}_x = \bar{K}_y =: \bar{K}_0$ for all x, y, and so that $\delta(Q_{x,\bar{K}_0+m}, \mathbb{R}^d) = mA \pm \varepsilon$ for $m \geq 1$. For simplicity, we will assume that our lattice fulfils these properties.

If $\int_{B(x,1)} |y-x|^{-n}\, d\mu(y) = \int_{\mathbb{R}^d \setminus B(x,1)} |y-x|^{-n}\, d\mu(y) = \infty$, then all the cubes $Q_{x,k}$, $k \in \mathbb{Z}$, satisfy $0 < \ell(Q_{x,k}) < \infty$. That is, they are transit cubes.

We denote $\mathcal{D}_k = \{Q_{x,k} : x \in \operatorname{supp}(\mu)\}$ for $k \in \mathbb{Z}$, and $\mathcal{D} = \bigcup_{k\in\mathbb{Z}} \mathcal{D}_k$.

Consider a cube $Q \subset \mathbb{R}^d$ whose center may not be in $\operatorname{supp}(\mu)$. Let $Q_{x,k}$ be one of the smallest cubes in $\mathcal{D}$ containing Q in the following sense. Set

$$\ell = \inf\{\ell(Q_{x,j}) : Q_{x,j} \in \mathcal{D},\, Q \subset Q_{x,j}\}.$$

Take $Q_{x,k}$ containing Q such that $\ell(Q_{x,k}) \leq \frac{100}{99}\ell$. Then we write $Q \in \mathcal{AD}_k$ (by the property (e) in Lemma 3.3, k depends only on Q). In a sense, Q is approximately in $\mathcal{D}_k$. Given k, j with $-\infty \leq k \leq j \leq +\infty$, we also denote $\mathcal{AD}_{k,j} = \bigcup_{h=k}^{j} \mathcal{AD}_h$. If Q is such that there are cubes $Q_{x,k}$, $Q_{y,k-1}$ with $Q_{x,k} \subset Q \subset Q_{y,k-1}$, then it follows easily that $Q \in \mathcal{AD}_{k,k-1}$.

3.2. The kernels $s_k(x,y)$. For each $x \in \operatorname{supp}(\mu)$, $s_k(x,\cdot)$ is a non negative radial non increasing function with center x, supported on $2Q_{x,k-1}$, and such that

(a) $s_k(x,y) \leq \dfrac{1}{A|x-y|^n}$ for all $y \in \mathbb{R}^d$.

(b) $s_k(x,y) \approx \dfrac{1}{A\ell(Q_{x,k})^n}$ for all $y \in Q_{x,k}$.

(c) $s_k(x,y) = \dfrac{1}{A|x-y|^n}$ for all $y \in Q_{x,k-1} \setminus Q_{x,k}$.

(d) $\nabla_y s_k(x,y) \leq C\, A^{-1} \min\left(\dfrac{1}{\ell(Q_{x,k})^{n+1}}, \dfrac{1}{|x-y|^{n+1}}\right)$ for all $y \in \mathbb{R}^d$.

Lemma 3.5. *If* $y \in \operatorname{supp}(\mu)$*, then* $\operatorname{supp}(s_k(\cdot, y)) \subset Q_{y,k-2}$*. If* $Q \in \mathcal{AD}_k$ *and* $z \in Q \cap \operatorname{supp}(\mu)$*, then* $\operatorname{supp}(s_{k+m}(z, \cdot)) \subset \frac{11}{10}Q$ *for all* $m \geq 3$*, and* $\operatorname{supp}(s_{k+m}(\cdot, z)) \subset \frac{11}{10}Q$ *for all* $m \geq 4$*.*

Proof: For the assertion $\operatorname{supp}(s_k(\cdot, y)) \subset Q_{y,k-2}$, see [**To3**] or [**To4**].

Let $Q \in \mathcal{AD}_k$ and $z \in Q \cap \operatorname{supp} \mu$. We have $Q \not\subset Q_{z,k+1}$, because otherwise $Q \notin \mathcal{AD}_k$. Thus $\ell(Q_{z,k+1}) \leq 2\ell(Q)$. Then,

$$\operatorname{supp}(s_{k+m}(z, \cdot)) \subset 2Q_{z,k+m-1} \subset \frac{11}{10}Q,$$

because $\ell(2Q_{z,k+m-1}) \leq \frac{2}{100}\ell(Q_{z,k+1}) \leq \frac{4}{100}\ell(Q)$. Finally, the inclusion $\operatorname{supp}(s_{k+m}(\cdot, z)) \subset \frac{11}{10}Q$ follows in a similar way. □

In [**To4**, Section 3] the following estimates are proved.

Lemma 3.6. *If A is big enough, then for all* $k \in \mathbb{Z}$ *and* $z \in \operatorname{supp}(\mu)$ *we have*

$$\int s_k(z, y)\, d\mu(y) \leq \frac{10}{9} \quad \text{and} \quad \int s_k(x, z)\, d\mu(x) \leq \frac{10}{9}. \tag{3.1}$$

If moreover $Q_{z,k}$ *is a transit cube, then*

$$\int s_k(z, y)\, d\mu(y) \geq \frac{9}{10} \quad \text{and} \quad \int s_k(x, z)\, d\mu(x) \geq \frac{9}{10}. \tag{3.2}$$

In the following lemma we state another technical result that we will need.

Lemma 3.7. *For all* $k \in \mathbb{Z}$ *and* $x, y \in \operatorname{supp}(\mu)$*, we have*

$$s_k(x, y) \leq C\big(s_{k-1}(y, x) + s_k(y, x) + s_{k+1}(y, x)\big). \tag{3.3}$$

The proof follows easily from our construction. See also [**To3**, Lemma 7.8].

We will denote by S_k the integral operator associated with the kernel $s_k(x, y)$ and the measure μ. Observe that (3.1) implies that the operators S_k are bounded uniformly on $L^p(\mu)$, for all $p \in [1, \infty]$. Also, from (3.3) we get

$$S_k f(x) \leq C\big(S_{k-1}^* f(x) + S_k^* f(x) + S_{k+1}^* f(x)\big), \tag{3.4}$$

for $f \in L^1_{\mathrm{loc}}(\mu)$, $f \geq 0$, and $x \in \mathbb{R}^d$.

Notice that the only property in the definition of Calderón-Zygmund kernel which $s_k(x,y)$ may not satisfy is the gradient condition on the first variable. It is not difficult to check that if the functions $\ell(Q_{x,k})$ were Lipschitz (with respect to x) uniformly on k, then we would be able to define $\widetilde{s}_k(x,y)$ so that

$$|\nabla_x s_k(x,y)(y)| \leq \frac{C}{A|x-y|^{n+1}}, \tag{3.5}$$

in addition to the properties above. The following lemma solves this question.

Lemma 3.8. *The lattice $\mathcal{D}$ can be constructed so that the functions $\ell(Q_{x,k})$ are Lipschitz (with respect to $x \in \operatorname{supp}(\mu)$) uniformly on k and the properties (a)–(f) in Lemma 3.3 still hold. In this case, the operators S_k, $k \in \mathbb{Z}$, are CZO's with constants uniform on k.*

Proof: Suppose that the cubes $Q_{x,k} \in \mathcal{D}$ have already been chosen and the properties stated in Lemma 3.3 hold. Let us see how we can choose cubes $\widetilde{Q}_{x,k}$, substitutes of $Q_{x,k}$, such that $\psi_k(x) := \ell(\widetilde{Q}_{x,k})$ are Lipschitz functions on $\operatorname{supp}(\mu)$. For a fixed k, we set

$$\psi_k(x) := \sup_{z \in \operatorname{supp}(\mu)} (\ell(Q_{z,k}) - |x-z|). \tag{3.6}$$

It is easily seen that this is a non negative Lipschitz function, with constant independent of k. Then, we denote by $\widetilde{Q}_{x,k}$ the cube centered at x with side length $\psi_k(x)$.

We have to show that $\widetilde{Q}_{x,k}$ is a good choice as a cube of the scale k. Indeed, by (3.6) it is clear that $\ell(\widetilde{Q}_{x,k}) \geq \ell(Q_{x,k})$. Thus $Q_{x,k} \subset \widetilde{Q}_{x,k}$.

Take now $z_0 \in \operatorname{supp}(\mu)$ such that

$$\ell(Q_{z_0,k}) - |x-z_0| \geq \frac{99}{100} \ell(\widetilde{Q}_{x,k}).$$

We derive $|x-z_0| \leq \ell(Q_{z_0,k})$, and also $\ell(\widetilde{Q}_{x,k}) \leq 100\, \ell(Q_{z_0,k})/99$. Thus $x \in 2Q_{z_0,k}$ and $\widetilde{Q}_{x,k} \subset 4Q_{z_0,k}$. The inclusions $Q_{x,k} \subset \widetilde{Q}_{x,k} \subset 4Q_{z_0,k}$ imply $\delta(\widetilde{Q}_{x,k}, Q_{x,k}) \leq C_8 \ll \delta(\widetilde{Q}_{x,k}, Q_{x,k-1})$, with C_8 depending only on n, d, C_0. One can verify that the properties in Lemma 3.3 still hold, assuming that the constant C_8 is absorbed by the "error" ε in (c) and (d) in Lemma 3.3. □

3.3. The maximal operator N. In the following lemma we show which is the relationship between N and the operators S_k.

Lemma 3.9. *For all $f \in L^1_{\rm loc}(\mu)$, $x \in \mathbb{R}^d$, we have*
$$Nf(x) \approx \sup_{k\in\mathbb{Z}} S_k|f|(x),$$
with constants depending on A, C_0, n, d but independent of f and x.

Proof: For fixed $x \in \operatorname{supp}(\mu)$ and $k \in \mathbb{Z}$, we have $s_k(x,y) \le C\,\varphi_{x,r,R}(y)$, with $r = C\ell(Q_{x,k})$ and $R = C\ell(Q_{x,k-1})$. Assume $0 < r, R < \infty$. Since $\|\varphi_{x,r,R}\|_{L^1(\mu)} \le C$ we get
$$S_k|f|(x) \le \frac{C}{1+\|\varphi_{x,r,R}\|_{L^1(\mu)}} \int |\varphi_{x,r,R} f|\, d\mu \le C\, Nf(x).$$
If $r = 0$ or $R = \infty$, we also have $S_k|f|(x) \le C\,Nf(x)$ by an approximation argument, and so $\sup_k S_k|f|(x) \le C\,Nf(x)$.

Let us see the converse inequality. Given $0 < r < R < \infty$, let k be the least integer such that $Q_{x,k} \subset B(x,r)$. Now let m be the least positive integer such that $B(x,R) \subset Q_{x,k-m}$. Then we have
$$\varphi_{x,r,R}(y) \le C\,(s_k(x,y) + s_{k-1}(x,y) + \cdots + s_{k-m}(x,y)).$$
Also, it is easily checked that $1 + \|\varphi_{x,r,R}\|_{L^1(\mu)} \ge C^{-1} m$. Therefore,
$$\frac{1}{1+\|\varphi_{x,r,R}\|_{L^1(\mu)}} \int |\varphi_{x,r,R} f|\, d\mu \le \frac{C}{m} \sum_{h=0}^{m} S_h|f|(x) \le C \sup_i S_i|f|(x). \qquad \square$$

In the rest of the paper we will assume that ***N* is defined not by (1.5), but as**
$$Nf(x) := \sup_{k\in\mathbb{Z}} S_k|f|(x).$$
With this new definition we have:

Lemma 3.10. *Let $\lambda > 0$ and $f \in L^1_{\rm loc}(\mu)$. For each $k \in \mathbb{Z}$, the set $\{x \in \mathbb{R}^d : S_k|f|(x) > \lambda\}$ is open. As a consequence, $\{x \in \mathbb{R}^d : Nf(x) > \lambda\}$ is open too.*

The proof is an easy exercise which is left for the reader.

Given a fixed $x \in \operatorname{supp}(\mu)$, we can think of $S_k f(x)$ as an average of the means $m_{B(x,r)} f := \int_{B(x,r)} f\, d\mu/\mu(B(x,r))$ over some range of radii r. Arguing in this way, (1.7) follows. We will exploit the same idea in the following lemma.

Lemma 3.11. *For all $\alpha > 1$, we can choose constants A, β, C_9 big enough so that the following property holds: Let $x \in \operatorname{supp}(\mu)$, $k \in \mathbb{Z}$ and $f \in L^1_{\rm loc}(\mu)$, and assume that $Q_{x,k}$ is a transit cube. Then there exists some ball $B(x,r)$ with $Q_{x,k} \subset B(x,\alpha^{-1}r)$, $B(x,r) \subset Q_{x,k-1}$ such that $B(x,\alpha^{-1}r)$ is (α,β)-doubling and $m_{B(x,r)}|f| \le C_9\, S_k|f|(x)$.*

It is easy to check that there are balls $B(x, r_1)$ and $B(x, r_2)$ with $Q_{x,k} \subset B(x, \alpha^{-1} r_1)$, $B(x, r_2) \subset Q_{x,k}$ such that $B(x, \alpha^{-1} r_1)$ is (α, β)-doubling and $m_{B(x,r_2)}|f| \leq C\, S_k|f|(x)$. However, it is more difficult to see that we may take $B(x, r_1) = B(x, r_2)$, as the lemma asserts.

On the other hand, the lemma is false if we substitute the condition "$m_{B(x,r)}|f| \leq C_9\, S_k|f|(x)$" by "$m_{B(x,r)}|f| \geq C_9^{-1}\, S_k|f|(x)$".

Proof: We denote $\lambda := S_k|f|(x)$, $R_0 = d^{1/2} \ell(Q_{x,k})$, and $R_1 = \ell(Q_{x,k-1})/2$. Recall that, for fixed x, k, we have defined $s_k(x, y) = \psi(|y - x|)$, where $\psi \colon \mathbb{R} \to \mathbb{R}$ is non negative, smooth, radial, and non increasing. Then,

$$\lambda = \int |f(y)| s_k(x, y)\, d\mu(y) = \int_0^\infty |\psi'(r)| \left(\int_{B(x,r)} |f|\, d\mu \right) dr.$$

We denote $h(r) = |\psi'(r)|\, \mu(B(x, \alpha^{-1} r))$ and

$$m_\alpha(r) = \frac{1}{\mu(B(x, \alpha^{-1} r))} \int_{B(x,r)} |f|\, d\mu.$$

Thus, $\lambda = \int_0^\infty h(r)\, m_\alpha(r)\, dr$.

Let us see that $\int_{\alpha R_0}^{R_1} h(r)\, dr$ is big. Recall that $\psi'(r) = 1/(A\, r^{n+1})$ for $r \in [R_0, R_1]$. Then, for $s \in [R_0, \alpha^{-1} R_1]$, we have $|\psi'(\alpha s)| = |\psi'(s)|/\alpha^{n+1}$. Therefore,

$$\begin{aligned}
\int_{\alpha R_0}^{R_1} h(r)\, dr &= \alpha^{-n} \int_{R_0}^{\alpha^{-1} R_1} |\psi'(s)|\, \mu(B(x, s))\, ds \\
&= \alpha^{-n} \Bigg(\int_{R_0 \leq |x-y| \leq \alpha^{-1} R_1} s_k(x, y)\, d\mu(y) \\
&\qquad + \psi(R_0)\, \mu(B(x, R_0)) - \psi(\alpha^{-1} R_1)\, \mu(B(x, \alpha^{-1} R_1)) \Bigg) \\
&\geq \alpha^{-n} \left(\int_{R_0 \leq |x-y| \leq \alpha^{-1} R_1} s_k(x, y)\, d\mu(y) - C_0\, A^{-1} \right).
\end{aligned}$$

Since $\int_{|x-y| \leq R_0} s_k(x, y)\, d\mu(y) \leq C\, A^{-1}$ and also $\int_{|x-y| \geq \alpha^{-1} R_1} s_k(x, y)\, d\mu(y) \leq C\, A^{-1}$, for A big enough we obtain

$$\int_{\alpha R_0}^{R_1} h(r)\, dr \geq \frac{1}{2\alpha^n} =: M,$$

using (3.2). If we denote the measure $h(r)\,dr$ by h, we get

$$h\{r \geq 0 : m_\alpha(r) > 2\lambda/M\} \leq \frac{M}{2\lambda} \int_0^\infty m_\alpha(r)\, h(r)\, dr = \frac{M}{2}.$$

Thus,

$$h\{r \in [\alpha R_0, R_1] : m_\alpha(r) \leq 2\lambda/M\} \geq M - \frac{M}{2} = \frac{M}{2}.$$

Now we will deal with the doubling property. If $B(x, \alpha^{-1} r)$ is not (α, β)-doubling, we write $r \in ND$. We have

$$\begin{aligned} h([\alpha R_0, R_1] \cap ND) &= \int_{r \in [\alpha R_0, R_1] \cap ND} |\psi'(r)|\, \mu(B(x, \alpha^{-1} r))\, dr \\ &\leq \beta^{-1} \int_{\alpha R_0}^{R_1} |\psi'(r)|\, \mu(B(x, r))\, dr \\ &\leq \beta^{-1} \int s_k(x, y)\, d\mu(y) \leq \frac{10}{9} \beta^{-1}. \end{aligned}$$

Therefore,

$$h\big(\{r \in [\alpha R_0,\, R_1] : m_\alpha(r) \leq 2\lambda/M\} \setminus ND\big) \geq \frac{M}{2} - \frac{10}{9}\beta^{-1}.$$

So if we take β big enough, there exists some $r \in [\alpha R_0, R_1]$ such that $B(x, \alpha^{-1} r)$ is (α, β)-doubling and $m_{B(x,r)}|f| \leq m_\alpha(r) \leq 2\lambda/M$. □

As a direct corollary of Lemma 3.11 we get:

Lemma 3.12. *Assume that A, β, C_{10} are positive and big enough. Let $x \in \operatorname{supp}(\mu)$, $k \in \mathbb{Z}$ and $f \in L^1_{\mathrm{loc}}(\mu)$. If $Q_{x,k}$ is a transit cube, then there exists some $(2, \beta)$-doubling cube $Q \in \mathcal{AD}_{k,k-1}$ centered at x such that $m_{2Q}|f| \leq C_{10}\, S_k|f|(x)$.*

In the rest of the paper we will assume that the constant A used to construct the lattice $\mathcal{D}$ and the kernels $s_k(x, y)$ has been chosen big enough so that the conclusion of the preceding lemma holds.

4. The main lemmas

Theorems 1.1, 1.2 and 1.3 follow from the following two lemmas:

Lemma 4.1. *Let p, γ be constants with $1 \leq p < \infty$ and $0 < \gamma \leq 1$. Let $w > 0$ be a weight and $\sigma = w^{-1/(p-1)}$ (for $p \neq 1$). The following statements are equivalent:*

(a) *All operators $T \in CZO(\gamma)$ are of weak type (p,p) with respect to $w\,d\mu$.*

(b) *For all $T \in CZO(\gamma)$, T_* is of weak type (p,p) with respect to $w\,d\mu$.*

(c) *The maximal operator N is of weak type (p,p) with respect to $w\,d\mu$.*

(d) *The operators S_k are of weak type (p,p) with respect to $w\,d\mu$ uniformly on $k \in \mathbb{Z}$.*

(e) *(Only in the case $p \neq 1$.) For all $k \in \mathbb{Z}$ and all cubes Q,*

$$\int |S_k(w\,\chi_Q)|^{p'}\sigma\,d\mu \leq C\,w(Q), \tag{4.1}$$

with C independent of k and Q.

Lemma 4.2. *Let p, γ be constants with $1 < p < \infty$ and $0 < \gamma \leq 1$. Let $w > 0$ be a weight and $\sigma = w^{-1/(p-1)}$. The following statements are equivalent:*

(a) *All operators $T \in CZO(\gamma)$ are bounded on $L^p(w)$.*

(b) *For all $T \in CZO(\gamma)$, T_* is bounded on $L^p(w)$.*

(c) *The maximal operator N is bounded on $L^p(w)$.*

(d) *The operators S_k are bounded on $L^p(w)$ uniformly on $k \in \mathbb{Z}$.*

(e) *For all $k \in \mathbb{Z}$ and all cubes Q,*

$$\int |S_k(\sigma\,\chi_Q)|^{p}w\,d\mu \leq C\,\sigma(Q)$$

and

$$\int |S_k(w\,\chi_Q)|^{p'}\sigma\,d\mu \leq C\,w(Q),$$

with C independent of k and Q.

Notice that the Sawyer type conditions (e) in Lemma 4.1 and Lemma 4.2 involve the operators S_k instead of the maximal operator N. In the present formulation these conditions are much weaker and of more geometric nature than the analogous conditions involving N.

The scheme for proving both lemmas is the same. In both cases we will start by (c) $\Rightarrow$ (b). Later we will see (b) $\Rightarrow$ (a) $\Rightarrow$ (d) $\Rightarrow$ (e). These will be the easy implications. Notice, for instance, that (b) $\Rightarrow$ (a) is trivial. Finally we will show (e) $\Rightarrow$ (c) (except in the weak $(1,1)$ case). This will be the most difficult part of the proof (in both lemmas). In the weak $(1,1)$ case, we will see directly the implication (d) $\Rightarrow$ (c).

For simplicity, to prove Lemmas 4.1 and 4.2, we will assume that **all the cubes $Q_{x,k} \in \mathcal{D}$ are transit cubes**. In Section 11 we will give some hints for the proof in the general case. We have operated in this way because the presence of stopping cubes in the lattice $\mathcal{D}$ introduces some technical difficulties which make the proofs more lengthy, but the ideas and arguments involved are basically the same than in the special case in which all the cubes in $\mathcal{D}$ are transit cubes.

First we will prove Lemma 4.1.

5. The implication (c) $\Rightarrow$ (b) of Lemma 4.1

Definition 5.1. We say that w satisfies the Z_∞ property if there exists some constant $\tau > 0$ such that for any cube $Q \in \mathcal{AD}_k$ and any set $A \subset \mathbb{R}^d$, if

$$S_{k+3}\chi_A(x) \geq 1/4 \quad \text{for all } x \in Q, \tag{5.1}$$

then $w(A \cap 2Q) \geq \tau\, w(Q)$.

Lemma 5.2. *If*

$$\int |S_k^*(w\,\chi_Q)|^{p'}\,\sigma\,d\mu \leq C\,w(Q)$$

for all cubes Q and all $k \in \mathbb{Z}$, then w satisfies the property Z_∞.

Proof: Take $Q \in \mathcal{AD}_k$ and $A \subset \mathbb{R}^d$ satisfying (5.1). By the assumption above, the fact that $\operatorname{supp}(s_{k+3}(x,\cdot)) \subset 2Q$ for $x \in Q$, and Hölder's inequality, we get

$$\begin{aligned} w(Q) &\leq 4 \int_Q (S_{k+3}\chi_A)\,w\,d\mu \\ &= 4 \int (S_{k+3}\chi_{A\cap 2Q})\,w\,d\mu \\ &= 4 \int_{A\cap 2Q} S_{k+3}^*(w\,\chi_Q)\,d\mu \\ &\leq 4 \left(\int S_{k+3}^*(w\,\chi_Q)^{p'}\,\sigma\,d\mu\right)^{1/p'} w(A\cap 2Q)^{1/p} \\ &\leq C\,w(Q)^{1/p'}\,w(A\cap 2Q)^{1/p}, \end{aligned}$$

and so $w(Q) \leq C\,w(A \cap 2Q)$. $\square$

Notice that if N is bounded on $L^r(w)$ for some $r \in (1,\infty)$ or of weak type (r,r), then the operators S_k are bounded on $L^p(w)$ uniformly on k for any p with $r < p < \infty$. By duality, the operators S_k^* are bounded on $L^{p'}(\sigma)$ uniformly on k too. Then, by Lemma 5.2, w satisfies Z_∞.

Occasionally we will apply the Z_∞ condition by means of the following lemma.

Lemma 5.3. *Suppose that w satisfies the Z_∞ property. Let $A \subset \mathbb{R}^d$ and $Q \in \mathcal{AD}_h$. Let $\{P_i\}_i$ be a family of cubes with finite overlap such that $A \cap \frac{3}{2}Q \subset \bigcup_i P_i$, with $P_i \in \mathcal{AD}_{+\infty,h+4}$ for all i. There exists some constant $\delta > 0$ such that if $\mu(A \cap P_i) \leq \delta\,\mu(P_i)$ for each i, then*

$$w(2Q \setminus A) \geq \tau\, w(Q), \tag{5.2}$$

for some constant $\tau > 0$ (depending on Z_∞). If, moreover, $w(2Q) \leq C_{11}\, w(Q)$, then

$$w(A \cap 2Q) \leq (1 - C_{11}^{-1}\tau)\, w(2Q). \tag{5.3}$$

This lemma, specially the inequality (5.3), shows that the Z_∞ property can be considered as a weak version of the usual A_∞ property satisfied by the A_p weights. Notice that unlike A_∞, the Z_∞ condition is not symmetric on μ and w.

Let us remark that we have not been able to prove that the constant $1 - C_{11}^{-1}\tau$ in (5.3) can be substituted by some constant C_δ tending to 0 as $\delta \to 0$. Many difficulties in the arguments below stem from this fact.

Proof: For a fixed $x_0 \in Q$, we denote $Q_0 := Q_{x_0,h+3}$. Observe that

$$\operatorname{supp}(s_{h+3}(x_0,\cdot)\,\chi_A) \subset A \cap 2Q_{x_0,h+2} \subset A \cap \frac{3}{2}Q \subset \bigcup_i P_i.$$

We have

$$S_{h+3}\chi_A(x_0) \leq C \sum_{j=1}^{n_0} \frac{\mu(2^j Q_0 \cap A)}{\ell(2^j Q_0)^n},$$

where n_0 is the least integer such that $2Q_{x_0,h+2} \subset 2^{n_0} Q_0$. If $P_i \cap 2^j Q_0 \neq \varnothing$, then $\ell(P_i) \leq \ell(2^j Q_0)/10$. Therefore,

$$\mu(2^j Q_0 \cap A) \leq \sum_{i: P_i \cap 2^j Q_0 \neq \varnothing} \mu(P_i \cap A) \leq \sum_{i: P_i \cap 2^j Q_0 \neq \varnothing} \delta\, \mu(P_i) \leq C\,\delta\,\mu(2^{j+1} Q_0).$$

Therefore,

$$S_{h+3}\chi_A(x_0) \leq C\,\delta \sum_{j=1}^{n_0} \frac{\mu(2^{j+1} Q_0)}{\ell(2^j Q_0)^n} \leq C\,\delta.$$

If δ is small enough, we have $S_{h+3}\chi_{\mathbb{R}^d\setminus A}(x_0) \geq 1/4$ for all $x_0 \in Q$, by (3.2). Thus (5.2) holds.

Finally, (5.3) follows easily from (5.2). $\square$

The implication (c) $\Rightarrow$ (b) of Lemma 4.1 is a direct consequence of the good λ inequality in next lemma.

Lemma 5.4. *Let $T \in CZO(\gamma)$ and w which satisfies the Z_∞ condition. There exists some $\eta > 0$ such that for all $\lambda, \varepsilon > 0$*

$$(5.4)\quad w\{x : T_*f(x) > (1+\varepsilon)\,\lambda,\ Nf(x) \leq \delta\lambda\} \leq (1-\eta)\,w\{x : T_*f(x) > \lambda\}$$

if $\delta = \delta(\eta, \varepsilon) > 0$ is small enough.

The constant δ depends also on the weak $(1,1)$ norm of T_* (with respect to μ) and on n, d, besides of η, ε, but not on λ.

Proof: Given $\lambda > 0$, we set $\Omega_\lambda = \{x : T_*f(x) > \lambda\}$ and

$$A_\lambda = \{x : T_*f(x) > (1+\varepsilon)\,\lambda,\ Nf(x) \leq \delta\lambda\}.$$

So we have to see that there exists some $\eta > 0$ such that, for all $\varepsilon > 0$ and $\lambda > 0$, $w(A_\lambda) \leq (1-\eta)\,w(\Omega_\lambda)$ if we choose $\delta = \delta(\eta, \varepsilon) > 0$ small enough.

Since Ω_λ is open, we can consider a Whitney decomposition of it. That is, we set $\Omega_\lambda = \bigcup_i Q_i$, so that the cubes Q_i have disjoint interiors, $\operatorname{dist}(Q_i, \mathbb{R}^d \setminus \Omega_\lambda) \approx \ell(Q_i)$ for each i, and the cubes $4Q_i$ have finite overlap.

Take a cube Q_i such that there exists some $x_0 \in 2Q_i$ with $Nf(x_0) \leq \delta\lambda$. By standard arguments, one can check that for any $x \in 2Q_i$,

$$T_*(f\chi_{\mathbb{R}^d\setminus 3Q_i})(x) \leq \lambda + C\,M_R^c f(x),$$

where M_R^c is the centered radial maximal Hardy-Littlewood operator:

$$M_R^c f(x) = \sup_{r>0} \frac{1}{r^n} \int_{B(x,r)} |f|\,d\mu.$$

Since $M_R^c f \leq C\,Nf$, we get $T_*f\chi_{\mathbb{R}^d\setminus 3Q_i}(x) \leq (1+C\,\delta)\,\lambda$ if $x \in A_\lambda \cap 2Q_i$. For δ small enough, this implies $T_*(f\chi_{3Q_i})(x) \geq \frac{\varepsilon}{2}\,\lambda$ for all $x \in A_\lambda \cap 2Q_i$. So we have

$$(5.5)\qquad A_\lambda \cap 2Q_i \subset \{x \in 2Q_i : T_*(f\chi_{3Q_i})(x) > \varepsilon\lambda/2,\ Nf(x) \leq \delta\lambda\}.$$

Let $k \in \mathbb{Z}$ be such that $Q_i \in \mathcal{AD}_k$. Let us check that

$$(5.6)\qquad S_{k+3}\chi_{A_\lambda}(y) \leq C\,\delta\varepsilon^{-1} \quad \text{for all } y \in Q_i.$$

For a fixed $y \in Q_i$, let j_0 be the least non negative integer such that there exists some $y_0 \in 2^{j_0}Q_{y,k+3} \cap A_\lambda$. Let us denote $C_j = 2^j Q_{y,k+3} \setminus 2^{j-1}Q_{y,k+3}$ for $j > j_0$, and $C_{j_0} = 2^{j_0}Q_{y,k+3}$. Then we have

$$S_{k+3}\chi_{A_\lambda}(y) = \int_{A_\lambda \cap 2Q_i} s_{k+3}(y,z)\,d\mu(z) \leq C \sum_{j=j_0}^{n_0} \frac{\mu(A_\lambda \cap C_j)}{\ell(2^j Q_{y,k+3})^n},$$

where n_0 is the least integer such that $2Q_{y,k+2} \subset 2^{n_0}Q_{y,k+3}$. Let V_j be the $\ell(2^jQ_{y,k+3})$-neighborhood of C_j. It is easily checked that $T_*(f\chi_{3Q_i \setminus V_j})(z) \leq C\,Nf(z)$ for all $z \in C_j$. Therefore, if δ is small enough, for $z \in A_\lambda \cap C_j$ we must have $T_*(f\chi_{V_j})(z) \geq \varepsilon\lambda/4$. Then, by the weak $(1,1)$ boundedness of T_* with respect to μ, we get

$$\mu(A_\lambda \cap C_j) \leq \mu\{z : T_*(f\chi_{V_j})(z) \geq \varepsilon\lambda/4\} \leq \frac{C}{\varepsilon\lambda}\int_{V_j} |f|\,d\mu.$$

Using the finite overlap of the neighborhoods V_j,

$$S_{k+3}\chi_{A_\lambda}(y) \leq \frac{C}{\varepsilon\lambda}\sum_{j=j_0}^{n_0} \frac{1}{\ell(2^jQ_{y,k+3})^n}\int_{V_j}|f|\,d\mu \leq \frac{C}{\varepsilon\lambda}Nf(y_0) \leq \frac{C\delta}{\varepsilon},$$

which proves (5.6).

By (3.2), we get $S_{k+3}\chi_{\mathbb{R}^d \setminus A_\lambda}(y) \geq 9/10 - C\delta\varepsilon^{-1} > 1/4$ for all $y \in Q_i$, if δ is small enough. By the Z_∞ condition, $w(2Q_i \setminus A_\lambda) \geq \tau w(Q_i)$. Therefore, by the finite overlap of the cubes $2Q_i$,

$$w(\Omega_\lambda) \leq \tau^{-1}\sum_i w(2Q_i \setminus A_\lambda) \leq C\,\alpha^{-1}w(\Omega_\lambda \setminus A_\lambda). \tag{5.7}$$

Thus, $w(A_\lambda) \leq (1 - C\,\tau^{-1})\,w(\Omega_\lambda)$. Now we only have to take $\eta := 1 - C\,\tau^{-1}$ (which does not depend on δ, ε or λ), and (5.4) follows. □

6. The implications (b) ⇒ (a) ⇒ (d) ⇒ (e) of Lemma 4.1

The implication (b) ⇒ (a) is trivial. Let us see the remaining ones.

Proof of (a) ⇒ (d) in Lemma 4.1: We have defined the kernels $s_k(x,y)$ so that they are CZ kernels uniformly on $k \in \mathbb{Z}$. By the statement (a) in Lemma 4.1 we know that they are of weak type (p,p) with respect to $w\,d\mu$. We only have to check that this holds *uniformly* on k. Indeed, if this is not the case, for each $m \geq 1$ we take S_{k_m} such that $\|S_{k_m}\|_{L^p(w),\,L^{p,\infty}(w)} \geq m^3$. Then we define $T = \sum_{m=1}^\infty \frac{1}{m^2}S_{k_m}$. Since $\sum_m \frac{1}{m^2} < \infty$, T is a CZO (using also uniform estimates for the operators S_k). On the other hand, we have $\|T\|_{L^p(w),\,L^{p,\infty}(w)} \geq$

$\frac{1}{m^2}\|S_{k_m}\|_{L^p(w),\,L^{p,\infty}(w)} \geq m$ for each m, because S_k are integral operators with non negative kernel. Thus $\|T\|_{L^p(w),\,L^{p,\infty}(w)} = \infty$, which contradicts the statement (a) in Lemma 4.1. □

Proof of (d) $\Rightarrow$ (e) in Lemma 4.1 for $1 < p < \infty$: Since the operators S_k are of weak type (p,p) with respect to $w\,d\mu$, from (3.4) it follows that their duals are also of weak type (p,p) with respect to $w\,d\mu$, uniformly on k. Then, the statement (e) is a consequence of duality in Lorentz spaces. We only have to argue as in [**Saw1**, p. 341], for example:

$$\begin{aligned}\left(\int |S_k(w\,\chi_Q)|^{p'}\sigma\,d\mu\right)^{1/p'} &= \sup_{\|f\|_{L^p(\sigma)}\leq 1}\int S_k(w\,\chi_Q)\,f\,\sigma\,d\mu \\ &= \sup_{\|f\|_{L^p(\sigma)}\leq 1}\int_Q S_k^*(f\,\sigma)\,w\,d\mu \\ &= \sup_{\|f\|_{L^p(\sigma)}\leq 1}\int_0^\infty w\{x\in Q : S_k^*(f\,\sigma)(x)>\lambda\}\,d\lambda \\ &\leq \int_0^\infty \min\big(C\,\lambda^{-p},\, w(Q)\big)\,d\lambda = C\,w(Q)^{1/p'}. \quad \square\end{aligned}$$

7. The implication (e) $\Rightarrow$ (c) of Lemma 4.1

We need to introduce some notation and terminology. Let Ω be an open set. Suppose that we have a Whitney decomposition $\Omega = \bigcup_i Q_i$ into dyadic cubes Q_i with disjoint interiors, with $10Q_i \subset \Omega$, $\operatorname{dist}(Q_i, \partial\Omega) \approx \ell(Q_i)$, and such that the cubes $4Q_i$ have finite overlap. We say that two cubes Q and R are neighbors if $Q \cap R \neq \varnothing$ (recall that we are assuming that the cubes are closed). For a fixed i, we denote by $U_1(Q_i)$ the union of all the neighbors of $3Q_i$ (including Q_i itself). For $m > 1$, inductively we let $U_m(Q_i)$ be the union of all the cubes which are neighbors of some cube in $U_{m-1}(Q_i)$. That is, one should think that $U_m(Q_i)$ is formed by $3Q_i$ and m "layers" of neighbors.

We denote by M_R the non centered radial maximal Hardy-Littlewood operator:

$$M_R f(x) = \sup_{B:x\in B}\frac{1}{r(B)^n}\int_B |f|\,d\mu,$$

where B stands for any ball containing x and $r(B)$ is its radius.

In order to prove the implication (e) $\Rightarrow$ (c) we will need a very sharp maximum principle. In the following lemma we deal with this question.

Lemma 7.1. *Let $\varepsilon > 0$ be some arbitrary fixed constant. There exist $\beta > 0$ and $m \geq 1$, $m \in \mathbb{Z}$, both big enough, such that the operator $T = N + \beta\, M_R$ satisfies the following maximum principle for all $\lambda > 0$ and all $f \in L^1_{\mathrm{loc}}(\mu)$: Let $\Omega_\lambda = \{x : Tf(x) > \lambda\}$, and consider a Whitney decomposition $\Omega_\lambda = \bigcup_i Q_i$ as above. Then, for any $x \in Q_i$,*

$$T(f\,\chi_{\mathbb{R}^d\setminus U_m(Q_i)})(x) \leq (1+\varepsilon)\,\lambda. \tag{7.1}$$

The point in this lemma is that the constant $\varepsilon > 0$ can be as small as we need, which will be very useful. We only have to define the operator T choosing β big enough, and also to take the integer m sufficiently big in $U_m(Q_i)$. Notice also that $Nf(x) \leq Tf(x) \leq (1+C\,\beta)Nf(x)$, because $M_Rf(x) \leq C\,Nf(x)$.

Proof: Let $x \in Q_i$ be some fixed point. First we will show that, for some $z \in \partial\Omega$,

$$M_R(f\,\chi_{\mathbb{R}^d\setminus U_m(Q_i)})(x) \leq (1+\varepsilon/2)\,M_Rf(z), \tag{7.2}$$

if we choose m big enough. Let B be some ball containing x such that

$$(1+\varepsilon/2)^{1/2}\frac{1}{r(B)^n}\int_B |f\,\chi_{\mathbb{R}^d\setminus U_m(Q_i)}|\,d\mu \geq M_R(f\,\chi_{\mathbb{R}^d\setminus U_m(Q_i)})(x).$$

Notice that if $M_R(f\,\chi_{\mathbb{R}^d\setminus U_m(Q_i)})(x) \neq 0$, then $B\setminus U_m(Q_i) \neq \varnothing$. Since $3Q_i \subset U_m(Q_i)$, we get

$$\operatorname{diam}(B) \geq \ell(Q_i). \tag{7.3}$$

Recall also that $U_m(Q_i)$ is formed by m "layers" of Whitney cubes, and so we have

$$\operatorname{diam}(B) \geq m \inf_{\substack{j: Q_j\subset U_m(Q_i)\\ Q_j\cap B\neq\varnothing}} \ell(Q_j). \tag{7.4}$$

We distinguish two cases:

(a) Assume $100\,\ell(Q_i) \leq [(1+\varepsilon/2)^{1/2n}-1]\,r(B) =: C_\varepsilon\, r(B)$. That is, $\ell(Q_i)$ is small compared to $r(B)$. We choose $z \in \partial\Omega$ such that $\operatorname{dist}(x,\mathbb{R}^d\setminus\Omega) = |x-z| \leq 100\ell(Q_i)$. Then there exists some ball B' containing z and B with radius $r(B') \leq r(B)+|x-z| \leq (1+\varepsilon/2)^{1/2n}\,r(B)$. Therefore,

$$M_Rf(z) \geq \frac{1}{r(B')^n}\int_{B'}|f|\,d\mu \geq \frac{1}{(1+\varepsilon/2)^{1/2}\,r(B)^n}\int_B |f|\,d\mu, \tag{7.5}$$

and (7.2) holds.

(b) Suppose that $100\,\ell(Q_i) \geq C_\varepsilon\, r(B)$. Then there exists some Whitney cube P in $U_m(Q_i)$ such that $P \cap B \neq \varnothing$ and $\ell(P) \leq 300\, C_\varepsilon^{-1} \ell(Q_i)/m$. Otherwise, by (7.4), $2\,r(B) \geq 300\, C_\varepsilon^{-1}\,\ell(Q_i)$, which contradicts our assumption.

Since $P \cap B \neq \varnothing$, we can find $z \in \partial\Omega$ such that $\operatorname{dist}(z, B) \leq 100\,\ell(P)$. Thus,

$$\operatorname{dist}(z, B) \leq 30000\, C_\varepsilon^{-1}\,\ell(Q_i)/m \leq C_\varepsilon\,\ell(Q_i)/2,$$

if m is chosen big enough. By (7.3), we obtain $\operatorname{dist}(z, B) \leq C_\varepsilon\, r(B)$. Then there exists some ball B' containing z and B with radius

$$r(B') \leq (1 + C_\varepsilon)\, r(B) = (1 + \varepsilon/2)^{1/2n}\, r(B).$$

Arguing as in (7.5), we obtain (7.2).

Now we have to deal with the term $Nf(x)$. Notice that if $z \in \partial\Omega$ is the point chosen in (a) or (b) above, then in both cases we have $|x - z| \leq C\,\ell(Q_i)$, where C may depend on m. Thus we may choose some constant $\eta > 0$ big enough so that $\eta\,\ell(Q_i) \gg \operatorname{dist}(x, \partial\Omega)$, $|x - z|$. We set $B_\eta := B(x, \eta\,\ell(Q_i))$, and we have

$$N(f\,\chi_{\mathbb{R}^d \setminus U_m(Q_i)})(x) \leq N(f\,\chi_{B_\eta \setminus U_m(Q_i)})(x) + N(f\,\chi_{\mathbb{R}^d \setminus B_\eta})(x).$$

Since $|x - z| \ll \eta\,\ell(Q_i)$, for each k we get

$$|S_k(f\,\chi_{\mathbb{R}^d \setminus B_\eta})(x) - S_k(f\,\chi_{\mathbb{R}^d \setminus B_\eta})(z)| \leq C_{12}\, M_R f(z),$$

where C_{12} may depend on η. Thus $N(f\,\chi_{\mathbb{R}^d \setminus B_\eta})(x) \leq Nf(z) + C_{12} M_R f(z)$. We also have $N(f\,\chi_{B_\eta \setminus U_m(Q_i)})(x) \leq C_{13}\, M_R f(z)$, with C_{13} depending on η. Therefore,

$$N(f\,\chi_{\mathbb{R}^d \setminus U_m(Q_i)})(x) \leq Nf(z) + C_\eta\, M_R f(z).$$

If we take β such that $C_\eta \leq \beta\varepsilon/2$, by (7.2), we obtain

$$\begin{aligned} T(f\,\chi_{\mathbb{R}^d \setminus U_m(Q_i)})(x) &\leq Nf(z) + C_\eta\, M_R f(z) + \beta\,(1 + \varepsilon/2)\, M_R f(z) \\ &\leq Nf(z) + \beta\,(1 + \varepsilon)\, M_R f(z) \\ &\leq (1 + \varepsilon)\, Tf(z) \leq (1 + \varepsilon)\,\lambda. \end{aligned}$$

□

Proof of (e) ⇒ (c) in Lemma 4.1 for $1 < p < \infty$*:* We will show that for some $\beta \geq 0$, the operator $T := N + \beta\, M_R$ is bounded on $L^p(w)$. The precise value of β will be fixed below. Without loss of generality, we take $f \in L^1(\mu)$ non negative with compact support. Given any $\lambda > 0$, we denote $\Omega_\lambda = \{x : Tf(x) > \lambda\}$. We will show that there exists some constant η, with $0 < \eta < 1$, such that for all $\varepsilon, \lambda > 0$

$$w(\Omega_{(1+\varepsilon)\lambda}) \leq \eta\, w(\Omega_\lambda) + \frac{C_\varepsilon}{\lambda^p} \int |f|^p w\, d\mu, \tag{7.6}$$

where C_ε is some constant depending on ε but not on λ. It is straightforward to check that (7.6) implies that T is of weak type (p,p) with respect to $w\, d\mu$ for ε small enough.

As in Lemma 7.1, we consider the Whitney decomposition $\Omega_\lambda = \bigcup_i Q_i$, where Q_i are dyadic cubes with disjoint interiors (the *Whitney cubes*). Suppose that m and β are chosen in Lemma 7.1 so that the maximum principle (7.1) holds with $\varepsilon/2$ instead of ε. Take some cube $Q_i \subset \Omega_\lambda$. To simplify notation, we will write U_i instead of $U_m(Q_i)$. Then, for $x \in Q_i \cap \Omega_{(1+\varepsilon)\lambda}$, we have $T(f\, \chi_{\mathbb{R}^d \setminus U_i})(x) \leq (1 + \varepsilon/2)\, \lambda$, and so

$$T(f\, \chi_{U_i})(x) \geq \varepsilon\, \lambda/2. \tag{7.7}$$

Let $h \in \mathbb{Z}$ be such that $Q_i \in \mathcal{AD}_h$. If for all k with $h - n_1 \leq k \leq h + 5$ we have $S_k(f\chi_{U_i})(x) \leq \delta\, \lambda$, where n_1, δ are positive constants which we will fix below, then we write $x \in B_\lambda$ (i.e. x is a "bad point") and, otherwise, $x \in G_\lambda$.

Notice that $G_\lambda \cup B_\lambda = \Omega_{(1+\varepsilon)\lambda} \subset \Omega_\lambda$. We will see that B_λ is quite small. Indeed, we will prove that

$$w(B_\lambda) \leq \eta_1 w(\Omega_\lambda), \tag{7.8}$$

for some positive constant $\eta_1 < 1$ independent of ε and λ.

Assume that (7.8) holds for the moment, and let us estimate $w(G_\lambda)$. For $Q_i \in \mathcal{AD}_h$, we have

$$\begin{aligned}
w(Q_i \cap G_\lambda) &\leq \frac{1}{\delta\lambda} \int_{Q_i} \sum_{k=h-n_1}^{h+5} S_k(f\, \chi_{U_i})\, w\, d\mu \\
&= \frac{1}{\delta\lambda} \sum_{k=h-n_1}^{h+5} \int f\, \chi_{U_i} S_k(w\, \chi_{Q_i})\, d\mu \\
&\leq \frac{1}{\delta\lambda} \sum_{k=h-n_1}^{h+5} \left(\int S_k(w\, \chi_{Q_i})^{p'}\, \sigma\, d\mu \right)^{1/p'} \left(\int_{U_i} |f|^p\, w\, d\mu \right)^{1/p} \\
&\leq \frac{C(n_1+6)}{\lambda} w(Q_i)^{1/p'} \left(\int_{U_i} |f|^p\, w\, d\mu \right)^{1/p}.
\end{aligned} \tag{7.9}$$

Using the inequality $a^{1/p'}\, b^{1/p} \leq \theta a + \theta^{-p'/p} b$, for $a, b, \theta > 0$, we get

$$w(Q_i \cap G_\lambda) \leq \theta\, w(Q_i) + \frac{C\theta^{-p'/p}}{\lambda^p} \int_{U_i} |f|^p\, w\, d\mu.$$

It is not difficult to check that the family of sets $\{U_i\}_i$ has bounded overlap (depending on m). Then, summing over all the indices i, we obtain

$$w(G_\lambda) \leq C\theta\, w(\Omega_\lambda) + \frac{C(\theta, m)}{\lambda^p} \int |f|^p\, w\, d\mu.$$

By (7.8), if we choose $\theta = (1 - \eta_1)/2C$, we get

$$w(\Omega_{(1+\varepsilon)\lambda}) \leq \frac{1 + \eta_1}{2} w(\Omega_\lambda) + \frac{C}{\lambda^p} \int |f|^p\, w\, d\mu,$$

which is (7.6) with $\eta = (1 + \eta_1)/2$.

Now we have to show that (7.8) holds. We intend to use the Z_∞ property. Let us see that

$$S_{h+3}\chi_{\mathbb{R}^d \setminus B_\lambda}(y) \geq \frac{1}{4} \tag{7.10}$$

for all $y \in Q_i$. By (7.7), if $z \in Q_i$, then $N(f\, \chi_{U_i})(z) \geq C_{14}\lambda$, where C_{14} is some positive constant depending on ε, β. Then we have

$$S_k(f\, \chi_{U_i})(z) \geq C_{14}\lambda \tag{7.11}$$

for some $k \geq h - n_1$. If moreover $z \in B_\lambda \cap Q_i$, then this inequality holds for some $k \geq h + 6$, assuming that we take $\delta < C_{14}$.

Suppose that $B_\lambda \cap \operatorname{supp}(s_{h+3}(y, \cdot)) \neq \varnothing$. Let $j_0 \geq 0$ be the least integer such there exists some $x_0 \in 2^{j_0} Q_{y,h+3}$, and let n_0 be the least integer such that $Q_{y,h+2} \subset 2^{n_0} Q_{y,h+3}$. Then we have

$$\begin{aligned} S_{h+3}\chi_{B_\lambda}(y) &= \int_{z \in B_\lambda} s_{h+3}(y, z)\, d\mu(z) \\ &\leq C \sum_{j=j_0}^{n_0} \frac{\mu(B_\lambda \cap (2^{j+1} Q_{y,h+3} \setminus 2^j Q_{y,h+3}))}{\ell(2^j Q_{y,h+3})^n}. \end{aligned}$$

It is not difficult to check that if $z \in B_\lambda \cap (2^{j+1} Q_{y,h+3} \setminus 2^j Q_{y,h+3})$ and $k \geq h + 6$, then $\operatorname{supp}(s_k(z, \cdot)) \subset 2^{j+2} Q_{y,h+3} \setminus 2^{j-1} Q_{y,h+3} =: V_j$. Therefore, $N(f \chi_{V_j})(z) \geq C_{14}\lambda$. Then, by the weak $(1,1)$ boundedness of N, we have

$$\begin{aligned} \mu(B_\lambda \cap (2^{j+1} Q_{y,h+3} \setminus 2^j Q_{y,h+3})) &\leq \mu\{z : N(f\, \chi_{V_j})(z) \geq C_{14}\lambda\} \\ &\leq \frac{C}{\lambda} \int_{V_j} |f|\, d\mu. \end{aligned}$$

Thus, by the finite overlap of the sets V_j, and since $x_0 \in B_\lambda$,

$$\begin{aligned} S_{h+3}\chi_{B_\lambda}(y) &\leq \frac{C}{\lambda} \sum_{j=j_0}^{n_0} \frac{1}{\ell(2^j Q_{y,h+3})^n} \int_{V_j} |f| \, d\mu \\ &\leq \frac{C}{\lambda}(S_{h+2}f(x_0) + S_{h+3}f(x_0) + S_{h+4}f(x_0)) \leq C_{15}\delta. \end{aligned}$$

Notice that C_{15} depends on ε, but not on δ. If δ is small enough, we obtain $S_{h+3}\chi_{B_\lambda}(y) \leq 1/4$. Now, we have $S_{h+3}\chi_{\mathbb{R}^d \setminus B_\lambda}(y) \geq 9/10 - S_{h+3}\chi_{B_\lambda}(y) \geq 1/4$, and (7.10) holds.

By the Z_∞ property, we get $w(2Q_i \setminus B_\lambda) \geq \tau \, w(Q_i)$, and because of the finite overlap of the cubes $2Q_i$,

$$w(\Omega_\lambda) \leq \tau^{-1} \sum_i w(2Q_i \setminus B_\lambda) \leq C_{16}\tau^{-1} w(\Omega_\lambda \setminus B_\lambda),$$

which implies (7.8). □

A slight modification of the arguments above yields the **proof of the implication (d) ⇒ (c) in the weak $(\mathbf{1},\mathbf{1})$ case**. Instead of using (4.1) to estimate $w(Q_i \cap G_\lambda)$ in (7.9), one can apply directly that the operators S_k are bounded from $L^1(w)$ into $L^{1,\infty}(w)$. We leave the details for the reader.

8. Preliminary lemmas for the proof of Lemma 4.2

Sections 8–10 are devoted to the proof of Lemma 4.2. As in our lemma about the weak (p,p) case, the implication (c) ⇒ (b) is a direct consequence of the good λ inequality of Lemma 5.4. The proofs of the implications (b) ⇒ (a) ⇒ (d) ⇒ (e) are similar to the ones of Lemma 4.1. We will not go through the details. So we have to concentrate on the implication (e) ⇒ (c), which is more difficult than the corresponding implication of the weak (p,p) case, as we will see.

In this section we will obtain some technical results which will be needed later for the proof of (e) ⇒ (c).

Lemma 8.1. *Let $\rho \geq 1$ be some fixed constant. Let $Q \subset \mathbb{R}^d$ be some cube and suppose that $x \in Q \cap \operatorname{supp}(\mu)$, $x' \in \rho Q \cap \operatorname{supp}(\mu)$, and $y \in \mathbb{R}^d \setminus 2Q$. Then, $s_k(x,y) \leq C_{17} \sum_{j=k-5}^{k+5} s_j(x',y)$, for any $k \in \mathbb{Z}$, with C_{17} depending on ρ and assuming A big enough (depending on ρ too).*

Proof: We denote $s^k(x',y) = \sum_{j=k-5}^{k+5} s_j(x',y)$. Observe that, by the definition of $s_j(x',y)$, we have

$$s^k(x',y) \approx \min\left(\frac{1}{A\ell(Q_{x',k+5})^n}, \frac{1}{A|x'-y|^n}\right) \quad \text{if } y \in Q_{x',k-5}. \tag{8.1}$$

Let $h \in \mathbb{Z}$ be such that $Q \in \mathcal{AD}_h$. If $k \geq h+3$, then $\operatorname{supp}(s_k(x,\cdot)) \subset 2Q$ by Lemma 3.5, and then $s_k(x,y) = 0$.

Assume now $k \leq h-3$. Since $Q \in \mathcal{AD}_h$, we have $Q \subset Q_{x,h-1}$. If A is big enough (depending on ρ), we deduce $x' \in Q_{x,h-2} \subset Q_{x,k}$ by (g) in Lemma 3.3. Then we get $2Q_{x,k-1} \subset Q_{x',k-4}$, and so $y \in Q_{x',k-4}$ if $s_k(x,y) \neq 0$. We also deduce $\ell(Q_{x',k+5}) \ll \ell(Q_{x,k})$. By (8.1), if $s_k(x,y) \neq 0$, we obtain

$$\begin{aligned} s^k(x',y) &\geq C^{-1} \min\left(\frac{1}{A\ell(Q_{x,k})^n}, \frac{1}{A|x'-y|^n}\right) \\ &\geq C^{-1} \min\left(\frac{1}{A\ell(Q_{x,k})^n}, \frac{1}{A|x-y|^n}\right) \geq s_k(x,y). \end{aligned}$$

Suppose finally that $|h-k| \leq 2$. As above, we have $x' \in Q_{x,h-2}$, and since $h-2 \geq k-4$, $x' \in Q_{x,k-4}$. Then we get $2Q_{x,k-1} \subset Q_{x',k-5}$, and so $y \in Q_{x',k-5}$ if $s_k(x,y) \neq 0$. On the other hand, $Q \not\subset Q_{x',h+1}$, because $Q \in \mathcal{AD}_h$. Thus

$$\ell(Q_{x',h+1}) \leq C\ell(Q) \leq C|x-y|,$$

with C depending on ρ. Then, if $s_k(x,y) \neq 0$, by (8.1) we get

$$s^k(x',y) \geq C^{-1} \frac{1}{A|x-y|^n} \geq C^{-1} s_k(x,y). \qquad \square$$

Given $\alpha, \beta > 1$, we say that some cube $Q \subset \mathbb{R}^d$ is μ-σ-(α,β)-doubling if $\mu(\alpha Q) \leq \beta\,\mu(Q)$ and $\sigma(\alpha Q) \leq \beta\,\sigma(Q)$. We say that Q is μ-σ-doubling if $\alpha = 2$ and β is some fixed constant big enough (which perhaps is not specified explicitly). Next lemma deals with the existence of this kind of cubes.

Lemma 8.2. *Suppose that the operators S_k are bounded on $L^r(\sigma)$ uniformly on k for some r with $1 < r < \infty$ and that the constant A is big enough. Then there exists some constant $\beta > 0$ such that for any $x \in \operatorname{supp}(\mu)$ and all cubes Q, R centered at x with $\delta(Q,R) \geq A/2$, there exists some μ-σ-$(100,\beta)$-doubling cube P centered at x, with $Q \subset P \subset R$.*

Proof: The constant β will be chosen below. For the moment, let us say that $\beta \geq 100^{d+1}$. Let N_0 be the least integer such that $R \subset 100^{N_0}Q$. For each $j \geq 0$, we denote $R_j = 100^{-j}R$. We have $\delta(R_{N_0},R) \geq A/2 -$

$100C_0 > A/4$. We will show that some cube R_j, with $j = 0, \ldots N_0$, is doubling with respect to μ and σ.

Suppose that all the cubes R_j, $j = 0, \ldots N_0$, are either non $(100, \beta)$-μ-doubling, or non $(100, \beta^r)$-σ-doubling (for simplicity, we will show the existence of a cube P which is $(100, \beta^r)$-σ-doubling, instead of $(100, \beta)$-σ-doubling). For each $j = 0, \ldots, N_0$, let a_j be the number of non $(100, \beta)$-μ-doubling cubes of the form $100^{-k}R$, $k = 0, \ldots, j$ and let b_j the number of non $(100, \beta^r)$-σ-doubling cubes of the form $100^{-k}R$, $k = 0, \ldots, j$. From our assumption we deduce

$$a_j + b_j \geq j + 1. \tag{8.2}$$

We have $\mu(R_j) \leq \beta^{-a_j}\mu(R)$. Thus,

$$\frac{\mu(R_j)}{\ell(R_j)^n} \leq \frac{\beta^{-a_j}\mu(R)}{100^{-jn}\ell(R)^n} \leq C_0 \frac{100^{jn}}{\beta^{a_j}}. \tag{8.3}$$

Let R_{s_0} be the largest non $(100, \beta^r)$-σ-doubling cube of the form $100^{-k}R$, $k = 0, \ldots, N_0$. Then, for $j \geq s_0$ we have

$$\sigma(R_j) \leq \beta^{-rb_j}\sigma(100R_{s_0}) \leq \frac{1}{2}\beta^{-rb_j}\sigma(100R_{s_0} \setminus R_{s_0}),$$

if β is big enough.

Let $h \in \mathbb{Z}$ be such that $Q \in \mathcal{AD}_h$. We denote $S = \sum_{i=-3}^{3} S_{h+i}$. From the properties of the kernels $s_k(x,y)$, it is easily seen that, for $x \in 100R_{s_0} \setminus R_{s_0}$ and $j = s_0, s_0+1, \ldots, N_0$, we have $S(\chi_{R_j})(x) \geq C^{-1}\mu(R_j)/\ell(R)^n$. Then, using the boundedness of S on $L^r(w)$, we obtain

$$\begin{aligned} C\,\sigma(R_j) &\geq \int_{100R_{s_0}\setminus R_{s_0}} |S(\chi_{R_j})|^r\,\sigma\,d\mu \geq C^{-1}\sigma(100R_{s_0} \setminus R_{s_0})\frac{\mu(R_j)^r}{\ell(R)^{nr}} \\ &\geq C^{-1}\frac{\beta^{b_j r}}{100^{jnr}}\sigma(R_j)\frac{\mu(R_j)^r}{\ell(R_j)^{nr}}, \end{aligned}$$

if $j \geq s_0$. Thus,

$$\frac{\mu(R_j)}{\ell(R_j)^n} \leq C\frac{100^{jn}}{\beta^{b_j}} \quad \text{if } j \geq s_0. \tag{8.4}$$

By (8.2), $\max(a_j, b_j) \geq (j+1)/2$. Then, from (8.3) and (8.4), we get

$$\frac{\mu(R_j)}{\ell(R_j)^n} \leq C\frac{100^{jn}}{\beta^{(j+1)/2}}.$$

Therefore,

$$\delta(R_{N_0}, R) \leq \sum_{j=0}^{\infty} C\left(\frac{100^n}{\beta^{1/2}}\right)^{j+1} \leq \frac{C}{\beta^{1/2}},$$

if $\beta^{1/2} > 2 \cdot 100^n$. Thus $\delta(R_{N_0}, R) \leq A/4$ if β is big enough, which is a contradiction. $\square$

Let us remark that if in the lemma above we also assume that the operators S_k are bounded uniformly on k on $L^s(w)$ for some s with $1 < s < \infty$, then it is possible to show the existence of cubes which are μ-doubling, σ-doubling and w-doubling simultaneously, by an easy modification of the proof.

Notice also that if $\int |S_k(\sigma\chi_Q)|^p \, w \, d\mu \leq C\,\sigma(Q)$ for $k \in \mathbb{Z}$ and all the cubes $Q \subset \mathbb{R}^d$, then N is of weak type (p', p') with respect to σ and bounded on $L^r(\sigma)$ for $p' < r \leq \infty$. Thus the assumptions of the preceding lemma are satisfied.

Lemma 8.3. *Suppose that the operators S_k are bounded on $L^r(\sigma)$ uniformly on k for some r with $1 < r < \infty$ and that the constant A is big enough. Then there exists some constant η with $0 < \eta < 1$ such that, for all $x \in \mathbb{R}^d$ and $k \in \mathbb{Z}$, $\sigma(Q_{x,k}) \leq \eta\sigma(Q_{x,k-1})$.*

Proof: We denote $S = \sum_{i=-2}^{2} S_{h+i}$. Then, we have $S(\chi_{Q_{x,k-1}\setminus Q_{x,k}})(y) \geq C_{18} > 0$, for all $y \in Q_{x,k}$. Therefore,

$$\sigma(Q_{x,k}) \leq C_{18}^{-r} \int |S(\chi_{Q_{x,k-1}\setminus Q_{x,k}})|^r \, \sigma \, d\mu \leq C_{19}\,\sigma(Q_{x,k-1} \setminus Q_{x,k}).$$

Thus, $\sigma(Q_{x,k-1}) \geq (1 + C_{19}^{-1})\,\sigma(Q_{x,k})$. $\square$

We will use the following version of Wiener's Covering Lemma.

Lemma 8.4. *Let $A \subset \mathbb{R}^d$ be a bounded set and $\{Q_i\}_{i\in I}$ some family of cubes such that $A \subset \bigcup_{i\in I} Q_i$, with $Q_i \cap A \neq \varnothing$ for each $i \in I$. Then there exists some finite or countable subfamily $\{Q_j\}_{j\in J}$, $J \subset I$, such that*

(1) $A \subset \bigcup_{j\in J} 20Q_j$.
(2) $2Q_j \cap 2Q_k = \varnothing$ *if* $j, k \in J$.
(3) *If* $j \in J$, $k \notin J$, *and* $2Q_j \cap 2Q_k \neq \varnothing$, *then* $\ell(Q_k) \leq 10\ell(Q_j)$.

The main novelty with respect to the usual Wiener's Lemma is the assertion (3). Although the lemma follows by standard arguments, for the sake of completeness we will show the detailed proof.

Proof: We will construct inductively the set $J = \{j_1, j_2, \dots\}$. Let $\ell_1 = \sup_{i\in I} \ell(Q_i)$. If $\ell_1 = \infty$, the lemma is straightforward. Otherwise, we take Q_{j_1} such that $\ell(Q_{j_1}) > d_1/2$. Assume that $Q_{j_1}, \dots, Q_{j_{m-1}}$ have been chosen. We set

$$\ell_m = \sup\Big\{\ell(Q_i) : 4Q_i \not\subset \textstyle\bigcup_{s=1}^{m-1} 20Q_{j_s}\Big\},$$

and we choose Q_{j_m} such that $\ell(Q_m) > \ell_m/2$ and $4Q_{j_m} \not\subset \bigcup_{s=1}^{m-1} 20Q_{j_s}$.

By construction, $A \subset \bigcup_{m=1}^{\infty} 20Q_{j_m}$, and also $\ell(Q_{j_m}) \geq \ell(Q_{j_s})/2$ for $s > m$. We have $2Q_{j_m} \cap 2Q_{j_s} = \varnothing$ for each $s = 1, \dots, m-1$, because otherwise $2Q_{j_m} \subset 10Q_{j_s}$, and then $4Q_{j_m} \subset 20Q_{j_s}$.

Finally we show that the third property holds. Suppose that $k \notin J$ and $2Q_{j_m} \cap 2Q_k \neq \varnothing$. If $\ell(Q_k) > 10\ell(Q_{j_m})$, it is easily seen that $4Q_{j_m} \subset 4Q_k$. Because of the definition of Q_{j_m}, we must have $4Q_k \subset \bigcup_{s=1}^{m-1} 20Q_{j_s}$ (otherwise $\ell_m \geq \ell(Q_k) > 10\ell(Q_{j_m})$, which is not possible). However the last inclusions imply $4Q_{j_m} \subset \bigcup_{s=1}^{m-1} 20Q_{j_s}$, which is a contradiction. $\square$

9. Boundedness of N over functions of type $\sigma\chi_Q$ on $L^p(w)$

The main result of this section is the following lemma.

Lemma 9.1. *If*

$$\int |S_k(\sigma\,\chi_Q)|^p\, w\, d\mu \leq C\,\sigma(Q)$$

for all cubes $Q \subset \mathbb{R}^d$ uniformly on $k \in \mathbb{Z}$, then

$$\int |N(\sigma\,\chi_Q)|^p\, w\, d\mu \leq C\,\sigma(\tfrac{11}{10}Q)$$

for all cubes $Q \subset \mathbb{R}^d$.

In a sense, Lemma 9.1 acts as a substitute of the usual reverse Hölder inequality for the classical A_p weights. Its proof will follow by a self improvement argument in the same spirit as the proof of the reverse Hölder inequality for the A_p weights.

Given $h \in \mathbb{Z}$ and $f \in L^1_{\text{loc}}(\mu)$, we denote

$$N^h f(x) = \sup_{k \geq h} S_k|f|(x).$$

The next technical result concentrates the main steps of the proof of Lemma 9.1.

Lemma 9.2. *Let $S = \sup_Q \sigma(\frac{11}{10}Q)^{-1} \int |N(\sigma\chi_Q)|^p\, w\, d\mu$, where the supremum is taken over all cubes $Q \subset \mathbb{R}^d$. Assume that*

$$\int |S_k(\sigma\,\chi_Q)|^p\, w\, d\mu \leq C\,\sigma(Q)$$

for all cubes $Q \subset \mathbb{R}^d$ uniformly on $k \in \mathbb{Z}$. Then, for all $\varepsilon > 0$, there exists some constant C_ε such that for any μ-σ-$(2,\beta)$-doubling cube $Q \in \mathcal{AD}_h$,

$$\int_Q |N^h(\sigma\,\chi_Q)|^p\, w\, d\mu \leq (C_\varepsilon + \varepsilon\, S)\,\sigma(Q). \tag{9.1}$$

Proof: The construction. Let Q_0 be some fixed μ-σ-$(2,\beta)$-doubling cube, with $Q_0 \in \mathcal{AD}_h$. We also denote $\lambda_0 := m_{Q_0}\sigma$. We will show that (9.1) holds for Q_0. To this end, by an inductive argument, for each $k \geq 1$ we will construct a sequence of μ-σ-doubling cubes $\{Q_i^k\}_i$.

First we will show how the cubes $\{Q_i^1\}$ are obtained. Let

$$\Omega_0 = \{N^{h+20}\sigma(x) > K\lambda_0\}, \tag{9.2}$$

where K is some big positive constant which will be fixed below. By Lemma 3.10, this set is open. We consider some Whitney decomposition $\Omega_0 = \bigcup_i R_i^1$, where R_i^1 are dyadic cubes with disjoint interiors.

Let us check that $Q_0 \setminus \Omega_0 \neq \varnothing$. If $Q_0 \subset \Omega_0$, then for all $x \in Q_0 \cap \operatorname{supp}(\mu)$ there exists some cube Q_x centered at x, with $Q_x \in \mathcal{AD}_{+\infty,h+19}$ with $m_{Q_x}\sigma > C\,K\,\lambda_0$ (where $C > 0$ is some fixed constant). Since $Q_x \in \mathcal{AD}_{+\infty,h+19}$, we have $\ell(Q_x) \leq \ell(Q_0)/10$. By Besicovitch's Covering Theorem, there exists some covering $Q_0 \subset \bigcup_i Q_{x_i}$ with finite overlap. Using that Q_0 is σ-doubling, we obtain

$$\begin{aligned}\int_{Q_0} \sigma\,d\mu \geq C^{-1}\int_{2Q_0} \sigma\,d\mu \geq C^{-1}\sum_i \int_{Q_{x_i}} \sigma\,d\mu \\ \geq C^{-1}\,K\,\lambda_0 \sum_i \mu(Q_{x_i}) \geq C^{-1}\,K\,\lambda_0\,\mu(Q_0).\end{aligned}$$

Therefore, $m_{Q_0}\sigma \geq C^{-1}\,K\,\lambda_0$, which is a contradiction if K is big enough.

Since $Q_0 \setminus \Omega_0 \neq \varnothing$, by the properties of the Whitney covering, we have $\ell(R_j^1) \leq C_{20}\ell(Q_0)$ for any Whitney cube R_j^1 such that $R_j^1 \cap Q_0 \neq \varnothing$. Moreover, subdividing the Whitney cubes if necessary, we may assume that $C_{20} \leq 1/10$.

Let $g_j^1 \in \mathbb{Z}$ be such that $R_j^1 \in \mathcal{AD}_{g_j^1}$. Observe that if $R_j^1 \cap Q_0 \neq \varnothing$, then $R_j^1 \subset \frac{3}{2}Q_0$, and so $g_j^1 \geq h-2$. For $x \in R_j^1 \cap \operatorname{supp}(\mu)$, we consider some μ-σ-$(100,\beta)$-doubling cube $Q_x^1 \in \mathcal{AD}_{g_j^1+16}$, with β given by Lemma 8.2. Now we take a Besicovitch's covering of $Q_0 \cap \Omega_0$ with this type of cubes: $Q_0 \cap \Omega_0 \subset \bigcup_{i\in I_1} Q_i^1$, and we define $A_1 := \bigcup_{i\in I_1} Q_i^1$. Notice that, for each i, $10Q_i^1 \subset \frac{3}{2}Q_0$, because all the Whitney cubes intersecting Q_0 have side length $\leq \ell(Q_0)/10$. In particular, we have $A_1 \subset \frac{3}{2}Q_0$. For each $i \in I_1$, let $h_i^1 \in \mathbb{Z}$ be such that $Q_i^1 \in \mathcal{AD}_{h_i^1}$. If Q_i^1 is centered at some point in R_j^1, then $h_i^1 = g_j^1 + 16 \geq h + 14$. This finishes the step $k = 1$ of the construction.

Suppose now that the cubes $\{Q_i^k\}_{i\in I_k}$ (which are μ-σ-$(100,\beta)$-doubling, with $10Q_k^i \subset \frac{3}{2}Q_0$, and have finite overlap) have already been constructed. Let us see how the cubes $\{Q_i^{k+1}\}_{i\in I_{k+1}}$ are obtained. For each fixed cube Q_i^k we repeat the arguments applied to Q_0. We denote $\lambda_i^k = m_{Q_i^k}\sigma$ and let $h_i^k \in \mathbb{Z}$ be such that $Q_i^k \in \mathcal{AD}_{h_i^k}$. We consider the open set $\Omega_i^k = \{N^{h_i^k+20}\sigma(x) > K\lambda_i^k\}$, and a decomposition of it into Whitney dyadic cubes with disjoint interiors: $\Omega_i^k = \bigcup_j R_j^{k+1}$. Arguing as in the case of Q_0, we deduce $Q_i^k \setminus \Omega_i^k \neq \varnothing$, and if $R_j^k \cap Q_i^k \neq \varnothing$, then $\ell(R_j^k) \leq \ell(Q_i^k)/10$. Given $g_j^{k+1} \in \mathbb{Z}$ such that $R_j^{k+1} \in \mathcal{AD}_{g_j^{k+1}}$, for $x \in R_j^{k+1}$, we consider some μ-σ-$(100,\beta)$-doubling cube $Q_{i,x}^{k+1} \in \mathcal{AD}_{g_j^{k+1}+16}$.

It may happen that the union $\bigcup_{i\in I_k}(\Omega_i^k \cap Q_i^k)$ is not pairwise disjoint, and so for a fixed $x \in \bigcup_{i\in I_k}(\Omega_i^k \cap Q_i^k)$ there are several indices i such that $Q_{i,x}^{k+1}$ is defined. In any case, for each $x \in \bigcup_i(\Omega_i^k \cap Q_i^k)$ we choose $Q_x^{k+1} := Q_{i,x}^{k+1}$ with i so that $x \in \Omega_i^k \cap Q_i^k$ (no matter which i). Now we take a Besicovitch covering of $\bigcup_i(\Omega_i^k \cap Q_i^k)$ with cubes of the type Q_x^{k+1}. So we have $\bigcup_{i\in I_k}(\Omega_i^k \cap Q_i^k) \subset \bigcup_{j\in I_{k+1}} Q_j^{k+1}$, and the cubes Q_j^{k+1} have bounded overlap. Moreover, for each $j \in I_{k+1}$ there exists some i such that $10Q_j^{k+1} \subset \frac{3}{2}Q_i^k \subset \frac{3}{2}Q_0$. We define $A_{k+1} := \bigcup_{j\in I_{k+1}} Q_j^{k+1}$, and we denote by h_j^{k+1} the integer such that $Q_j^{k+1} \in \mathcal{AD}_{h_j^{k+1}}$.

The first step to estimate $\int_{Q_0} |N^h\sigma|^p\, w\, d\mu$. We want to show that given any $\varepsilon > 0$, if K is big enough, then

$$\int_{Q_0} |N^h(\sigma\chi_{Q_0})|^p\, w\, d\mu \leq (C_\varepsilon + \varepsilon\, S)\sum_{k=0}^{\infty} \sigma(A_k). \tag{9.3}$$

We will prove this estimate inductively. First we deal with the case $k = 0$. We have

$$\begin{aligned}\int_{Q_0} |N^h(\sigma\chi_{Q_0}|^p\, w\, d\mu &\leq \int_{Q_0} \sum_{k=h}^{h+19} |S_k(\sigma\chi_{Q_0})|^p\, w\, d\mu \\ &\quad + \int_{Q_0} |N^{h+20}\sigma|^p\, w\, d\mu \\ &\leq C\,\sigma(Q_0) + \int_{Q_0} |N^{h+20}\sigma|^p\, w\, d\mu.\end{aligned} \tag{9.4}$$

Given some small constant $\varepsilon > 0$, let $B_0 = \{x \in Q_0 : S_{h+3}\sigma(x) \leq \varepsilon\lambda_0\}$. Let us see that $\sigma(B_0)$ is small. By Lemma 3.12, for all $x \in B_0$ there exists some μ-doubling cube $P_x \in \mathcal{AD}_{+\infty,h+2}$ centered at x such that $m_{2P_x}\sigma \leq C\varepsilon\lambda_0$. We consider a Besicovitch's covering of B_0 with this type of cubes. That is, $B_0 \subset \bigcup_i P_{x_i}$, with $\sum_i \chi_{P_{x_i}} \leq C$. We have

$$\begin{aligned}\sum_i \sigma(2P_{x_i}) &\leq C\varepsilon\lambda_0 \sum_i \mu(2P_{x_i}) \leq C\varepsilon\lambda_0 \sum_i \mu(P_{x_i}) \\ &\leq C\varepsilon\lambda_0\, \mu(2Q_0) \leq C\varepsilon\lambda_0\, \mu(Q_0) = C\varepsilon\, \sigma(Q_0).\end{aligned}$$

In particular, we deduce $\sigma(B_0) \leq C\varepsilon\, \sigma(Q_0)$. Then we obtain

$$\begin{aligned}\int_{B_0} |N^{h+20}\sigma|^p\, w\, d\mu &\leq \sum_i \int_{P_{x_i}} |N^{h+20}\sigma|^p\, w\, d\mu \\ &= \sum_i \int_{P_{x_i}} |N^{h+20}(\sigma\chi_{\frac{3}{2}P_{x_i}})|^p\, w\, d\mu \\ &\leq S \sum_i \sigma(\tfrac{11}{10}\tfrac{3}{2}P_{x_i}) \\ &\leq S \sum_i \sigma(2P_{x_i}) \leq C\varepsilon S\, \sigma(Q_0).\end{aligned} \tag{9.5}$$

Now we have to estimate $\int_{Q_0 \setminus B_0} |N^{h+20}\sigma|^p\, w\, d\mu$. Given $x \in R_j^1 \subset \Omega_0$, let $x' \in \partial\Omega_0$ be such that $|x - x'| = \operatorname{dist}(x, \mathbb{R}^d \setminus \Omega)$. From Lemma 8.1, we derive the following maximum principle:

$$N^{h+25}(\sigma\chi_{\mathbb{R}^d \setminus 2R_j^1})(x) \leq C_{21} N^{h+20}\sigma(x') \leq C_{21} K\lambda_0, \tag{9.6}$$

where $C_{21} > 1$ is some fixed constant depending on C_0, n, d. Let us see that if $N^{h+25}\sigma(x) > 2C_{21}K\lambda_0$, then

$$N^{h+25}\sigma(x) \leq \max\Big(2 \max_{g_j^1 - 2 \leq t \leq g_j^1 + 4} S_t(\sigma\chi_{2R_j^1})(x),\ N^{g_j^1+5}(\sigma\chi_{2R_j^1})(x)\Big). \tag{9.7}$$

Indeed, we have

$$N^{h+25}\sigma(x) \leq \max\Big(\max_{h+25 \leq t \leq g_j^1+4} S_t\sigma(x),\ N^{g_j^1+5}\sigma(x)\Big),$$

(with equality if $h + 25 \leq g_j^1 + 5$). If $N^{h+25}\sigma(x) \leq N^{g_j^1+5}\sigma(x)$, then (9.7) follows from the fact that $N^{g_j^1+5}\sigma(x) = N^{g_j^1+5}(\sigma\chi_{2R_j^1})(x)$. If $N^{h+25}\sigma(x) = S_{t_0}\sigma(x)$ for some t_0 with $h + 25 \leq t_0 \leq g_j^1 + 4$, then $S_{t_0}\sigma(x) > 2C_{21}K\lambda_0$, and so

$$S_{t_0}(\sigma\chi_{2R_j^1})(x) \geq S_{t_0}\sigma(x) - N^{h+25}(\sigma\chi_{\mathbb{R}^d \setminus 2R_j^1})(x) \geq \frac{1}{2} S_{t_0}\sigma(x),$$

by (9.6). Thus,

$$N^{h+25}\sigma(x) \leq 2S_{t_0}(\sigma\chi_{2R_j^1})(x) \leq 2 \max_{h+25 \leq t \leq g_j^1+4} S_t(\sigma\chi_{2R_j^1})(x).$$

Moreover, it is easily checked that, for $t \leq g_j^1 - 2$ (and $x \in R_j^1$), we have $S_t(\sigma\chi_{2R_j^1})(x) \leq S_{g_j^1-2}(\sigma\chi_{2R_j^1})(x)$. Therefore, (9.7) holds in any case.

We denote $D_0 := \{x \in Q_0 : N^{h+25}\sigma(x) > 2C_{21}K\lambda_0\}$. Notice that $D_0 \subset \Omega_0 \cap Q_0 \subset A_1$. We have

$$\begin{aligned}\int_{Q_0\setminus B_0} |N^{h+20}\sigma|^p\, w\, d\mu \leq \sum_{t=h+20}^{h+24} \int_{Q_0\setminus B_0} |S_t\sigma|^p\, w\, d\mu \\ + \int_{Q_0\setminus B_0} |N^{h+25}\sigma|^p\, w\, d\mu \\ \leq C\,\sigma(Q_0) + \int_{Q_0\setminus B_0} |N^{h+25}\sigma|^p\, w\, d\mu,\end{aligned}$$

where we have used that $S_t\sigma(x) = S_t(\sigma\chi_{2Q_0})(x)$ if $t = 20, \dots, 24$ and $x \in Q_0$. Now we write

$$\int_{Q_0\setminus B_0} |N^{h+25}\sigma|^p\, w\, d\mu = \int_{Q_0\setminus(B_0\cup D_0)} + \int_{D_0\setminus B_0} =: I + II.$$

First we will estimate I. For $x \in Q_0 \setminus (B_0 \cup D_0)$, we have

$$N^{h+25}\sigma(x) \leq CK\lambda_0 \leq CK\varepsilon^{-1}\, S_{h+3}\sigma(x).$$

Therefore,

$$\begin{aligned}I = \int_{Q_0\setminus(B_0\cup D_0)} |N^{h+25}\sigma|^p\, w\, d\mu \leq CK^p\varepsilon^{-p} \int_{Q_0} |S_{h+3}\sigma|^p\, w\, d\mu \\ \leq CK^p\varepsilon^{-p}\sigma(2Q_0) \leq CK^p\varepsilon^{-p}\sigma(Q_0),\end{aligned}$$

where we have used that $S_{h+3}\sigma(x) = S_{h+3}(\sigma\chi_{2Q_0})(x)$.

It remains to estimate II. Given $x \in R_j^1 \cap (D_0 \setminus B_0)$, by (9.7) we get

$$\begin{aligned}\int_{R_j^1\cap(D_0\setminus B_0)} |N^{h+25}\sigma|^p\, w\, d\mu \leq C \sum_{t=g_j^1-2}^{g_j^1+39} \int |S_t(\sigma\chi_{2R_j^1})|^p\, w\, d\mu \\ + \int_{R_j^1\cap(D_0\setminus B_0)} |N^{g_j^1+40}\sigma|^p\, w\, d\mu \\ \leq C\,\sigma(2R_j^1) + \int_{R_j^1\cap(D_0\setminus B_0)} |N^{g_j^1+40}\sigma|^p\, w\, d\mu.\end{aligned}$$

Given $k \geq 1$, for $x \in A_k$, we denote $H_x^k := \max\{h_i^k : i \in I_k,\, x \in Q_i^k\}$. It is easily seen that if $x \in R_j^1 \cap A_1$, then $H_x^1 + 20 \leq g_j^1 + 40$. Then, summing over all the cubes $R_j^1 \subset \Omega_0$ such that $R_j^1 \cap Q_0 \neq \varnothing$, due to the finite overlap of the cubes $2R_j^1$, we obtain

$$\int_{D_0\setminus B_0} |N^{h+25}\sigma|^p \, w \, d\mu \leq C\,\sigma(2Q_0) + \int_{A_1} |N^{H_x^1+20}\sigma(x)|^p \, w(x)\, d\mu(x). \tag{9.8}$$

So we have shown that

$$\int_{Q_0} |N^h(\sigma\chi_{Q_0})|^p \, w\, d\mu \leq (C_{22}+C_{23}\varepsilon S)\,\sigma(Q_0) + \int_{A_1} |N^{H_x^1+20}\sigma|^p \, w\, d\mu, \tag{9.9}$$

with C_{22}, but not C_{23}, depending on K and ε.

The k-th step to estimate $\int_{Q_0} |N^h\sigma|^p \, w\, d\mu$. Now we will show that for any $k \geq 1$,

$$\begin{aligned}\int_{A_k} |N^{H_x^k+20}\sigma|^p \, w\, d\mu \leq (C'_{22} + C'_{23}\varepsilon S)\,\sigma(A_k) \\ + \int_{A_{k+1}} |N^{H_x^{k+1}+20}\sigma|^p \, w\, d\mu,\end{aligned} \tag{9.10}$$

with C'_{22}, but not C'_{23}, depending on K and ε. The arguments to prove (9.10) are similar to the ones we have used to obtain (9.9), although a little more involved because the cubes $\{Q_i^k\}_{i\in I_k}$ are non pairwise disjoint.

For each $i \in I_k$ we define $B_i^k = \{x \in Q_i^k : S_{h_i^k+3}\sigma(x) \leq \varepsilon\lambda_i^k\}$. Arguing as in (9.5), we deduce

$$\int_{B_i^k} |N^{h_i^k+20}\sigma|^p \, w\, d\mu \leq C\varepsilon S\,\sigma(Q_i^k).$$

We denote $B_k = \bigcup_{i\in I_k} B_i^k$. Using the definition of H_x^k, we obtain

$$\begin{aligned}\int_{B_k} |N^{H_x^k+20}\sigma(x)|^p \, w(x)\, d\mu(x) &\leq \sum_i \int_{B_i^k} |N^{H_x^k+20}\sigma(x)|^p \, w(x)\, d\mu(x) \\ &\leq \int_{B_i^k} |N^{h_i^k+20}\sigma|^p \, w\, d\mu \\ &\leq C\varepsilon S \sum_i \sigma(Q_i^k) \leq C\varepsilon S\sigma(A_k).\end{aligned} \tag{9.11}$$

To estimate $\int_{A_k\setminus B_k} |N^{H_x^k+20}\sigma|^p \, w\, d\mu$, we need to introduce some additional notation. Assume $I_k = \{1,2,3,\dots\}$. We denote $I_{k,t} := \{i \in I_k :$

$Q_i^k \in \mathcal{AD}_t\}$. We set

$$\widehat{Q}_i^k := Q_i^k \setminus \left(\bigcup_{l \in I_{k,t},\, t > h_i^k} Q_l^k \cup \bigcup_{l \in I_{k,h_i^k},\, l<i} Q_l^k \right).$$

It is easily checked that the sets $\widehat{Q}_i^k$, $i \in I_k$, are pairwise disjoint, that $\bigcup_{i\in I_k} \widehat{Q}_i^k = \bigcup_{i \in I_k} Q_i^k = A_k$, and moreover that if $x \in \widehat{Q}_i^k$, then $H_x^k = h_i^k$. We have

$$\begin{aligned}
\int_{A_k \setminus B_k} & |N^{H_x^k+20}\sigma|^p \, w \, d\mu \\
&= \sum_{i \in I_k} \int_{\widehat{Q}_i^k \setminus B_k} |N^{h_i^k+20}\sigma|^p \, w\, d\mu \\
&\leq \sum_{i\in I_k} \sum_{t=h_i^k+20}^{h_i^k+24} \int_{\widehat{Q}_i^k \setminus B_k} |S_t \sigma|^p \, w\, d\mu \\
&\quad + \sum_{i \in I_k} \int_{\widehat{Q}_i^k \setminus B_k} |N^{h_i^k+25}\sigma|^p \, w \, d\mu \\
&\leq C \sum_{i \in I_k} \sigma(2Q_i^k) + \int_{A_k \setminus B_k} |N^{H_x^k+25}\sigma|^p \, w\, d\mu \\
&\leq C\sigma(A_k) + \int_{A_k\setminus B_k} |N^{H_x^k+25}\sigma|^p \, w\, d\mu.
\end{aligned} \tag{9.12}$$

Now we set $D_i^k = \{x \in Q_i^k : N^{h_i^k+25}\sigma > 2C_{21}K\lambda_i^k\}$, and $D_k = \bigcup_{i \in I_k} D_i^k$. For $x \in Q_i^k \setminus (D_i^k \cup B_k)$, we have

$$N^{H_x^k+25}\sigma(x) \leq N^{h_i^k+25}\sigma(x) \leq CK\varepsilon^{-1}\, S_{h_i^k+3}\sigma(x).$$

Therefore, operating as in the case $k=0$, we get

$$\int_{Q_i^k \setminus D_i^k \cup B_k} |N^{H_x^k+25}\sigma(x)|^p \, w(x)\, d\mu(x) \leq CK^p \varepsilon^{-p} \sigma(Q_i^k).$$

Summing over $i \in I_k$, we obtain

$$\int_{A_k \setminus (B_k \cup D_k)} |N^{H_x^k+25}\sigma(x)|^p \, w(x)\, d\mu(x) \leq CK^p\varepsilon^{-p}\sigma(A_k). \tag{9.13}$$

Finally we deal with $\int_{D_k \setminus B_k} |N^{H_x^k+25}\sigma(x)|^p \, w(x)\, d\mu(x)$. For a fixed k, let $\{R_j^{k+1}\}_{j \in J_{k+1}}$ be the collection of *all* the Whitney cubes (originated from *all* the sets Ω_i^k, $i \in I_k$) such that if R_j^{k+1} comes from Ω_i^k, then $R_j^{k+1} \cap Q_i^k \neq \varnothing$. Assume $J_{k+1} = \{1, 2, 3, \dots\}$. We denote $J_{k+1,t} := \{j \in J_{k+1} : R_j^{k+1} \in \mathcal{AD}_t\}$. We set

$$\widehat{R}_j^{k+1} := \tfrac{3}{2} R_j^{k+1} \setminus \left(\bigcup_{l \in J_{k+1,t},\, t > g_j^{k+1}} \tfrac{3}{2} R_l^{k+1} \cup \bigcup_{l \in J_{k+1, g_j^{k+1}},\, l < j} \tfrac{3}{2} R_l^{k+1} \right).$$

The sets $\widehat{R}_j^{k+1}$, $j \in J_{k+1}$, are pairwise disjoint and

$$\bigcup_{j \in J_{k+1}} \widehat{R}_j^{k+1} = \bigcup_{j \in J_{k+1}} \frac{3}{2} R_j^{k+1} \supset A_{k+1}.$$

Moreover, it easily seen that if $x \in \widehat{R}_j^{k+1}$, then $g_j^{k+1} + 40 \geq H_x^{k+1} + 20$, and so $N^{g_j^{k+1}+40}\sigma(x) \leq N^{H_x^{k+1}+20}\sigma(x)$. If $\widehat{R}_j^{k+1}$ is originated by Ω_i^k, arguing as in the case $k = 0$, we deduce

$$\begin{aligned} &N^{H_x^k+25}\sigma(x) \leq N^{h_i^k+25}\sigma(x) \\ &\leq \max\Big(2 \max_{g_j^{k+1}-2 \leq t \leq g_j^{k+1}+39} S_t(\sigma\chi_{2R_j^{k+1}})(x),\, N^{g_j^{k+1}+40}(\sigma\chi_{2R_j^{k+1}})(x) \Big). \end{aligned}$$

Therefore,

$$\begin{aligned} &\int_{D_k \setminus B_k} |N^{H_x^k+25}\sigma(x)|^p \, w(x) \, d\mu(x) \\ &\qquad = \sum_{j \in J_{k+1}} \int_{\widehat{R}_j^{k+1} \cap (D_k \setminus B_k)} \\ &\qquad \leq \sum_{j \in J_{k+1}} \sum_{t = g_j^{k+1}-2}^{g_j^{k+1}+39} \int |S_t(\sigma\chi_{2R_j^{k+1}})|^p \, w \, d\mu \\ &\qquad\quad + \sum_{j \in J_{k+1}} \int_{\widehat{R}_j^{k+1} \cap (D_k \setminus B_k)} |N^{H_x^{k+1}+20}\sigma(x)|^p \, w(x) \, d\mu(x) \\ &\qquad \leq C \sum_{j \in J_{k+1}} \sigma(2R_j^{k+1}) + \int_{A_{k+1}} |N^{H_x^{k+1}+20}\sigma(x)|^p \, w(x) \, d\mu(x). \end{aligned} \tag{9.14}$$

In the following estimates the notation $R_j^{k+1} \sim Q_i^k$ means that R_j^{k+1} is a Whitney cube of Ω_i^k:

$$\begin{aligned}\sum_{j\in J_{k+1}} \sigma(2R_j^{k+1}) &= \sum_{i\in I_k} \sum_{j\in J_{k+1}: R_j^{k+1}\sim Q_i^k} \sigma(2R_j^{k+1}) \\ &\le C\sum_{i\in I_k} \sigma(\Omega_i^k \cap 2Q_i^k) \le C\sum_{i\in I_k}\sigma(Q_i^k) \le C\,\sigma(A_k).\end{aligned} \tag{9.15}$$

By (9.11), (9.12), (9.13), (9.14) and (9.15), (9.10) follows.

From (9.9) and (9.10), we get

$$\begin{aligned}\int_{Q_0} N(\sigma\chi_{Q_0})|^p\, w\, d\mu \le (C + C\varepsilon S)\sum_{k=0}^{\infty}\sigma(A_k) \\ + \limsup_{k\to\infty}\int_{A_k} |N^{H_x^k+20}\sigma(x)|^p\, w(x)\, d\mu(x).\end{aligned} \tag{9.16}$$

This is the same as (9.3), except for the last term on right hand side. However, we will see below that this term equals 0.

The estimate of $\sum_k \sigma(A_k)$. We are going to prove that

$$\sum_{k=0}^{\infty}\sigma(A_k) \le C\,\sigma(Q_0). \tag{9.17}$$

We denote $\widetilde{A}_k = \bigcup_{i\in I_k} 2Q_i^k$. It is easily seen that $\widetilde{A}_{k+1} \subset \widetilde{A}_k$ for all k (this is the main advantage of $\widetilde{A}_k$ over A_k). We will show that there exists some positive constant $\tau_0 < 1$ such that

$$\sigma(\widetilde{A}_{k+2}) \le \tau_0\,\sigma(\widetilde{A}_k) \tag{9.18}$$

for all k. This implies (9.17), because $\widetilde{A}_0, \widetilde{A}_1 \subset 2Q_0$ and Q_0 is σ-doubling.

For a fixed $k \ge 1$, by the covering Lemma 8.4, there exists some subfamily $\{2Q_j^k\}_{j\in I_k^0} \subset \{2Q_i^k\}_{i\in I_k}$ such that

(1) $\widetilde{A}_k \subset \bigcup_{j\in I_k^0} 40Q_j^k$.

(2) $4Q_j^k \cap 4Q_l^k = \varnothing$ if $j, l \in I_k^0$.

(3) If $j \in I_k^0$, $l \notin I_k^0$, and $4Q_j^k \cap 4Q_l^k \ne \varnothing$, then $\ell(Q_l^k) \le 10\ell(Q_j^k)$.

First, we will see that

$$\sigma(2Q_j^k \cap \widetilde{A}_{k+2}) \leq \tau_1\, \sigma(2Q_j^k) \quad \text{if } j \in I_k^0, \tag{9.19}$$

for some fixed constant $0 < \tau_1 < 1$. By Lemma 5.3 it is enough to show that, for each $x \in \frac{3}{2}Q_j^k \cap \operatorname{supp}(\mu)$, there exists some cube $P \in \mathcal{AD}_{h_j^k+4}$ centered at x such that $\mu(\widetilde{A}_{k+2} \cap P) \leq \delta_0\, \mu(P)$, with δ_0 sufficiently small.

Let $2Q_s^{k+1}$ some cube which forms $\widetilde{A}_{k+1}$ such that $2Q_s^{k+1} \cap 2Q_j^k \neq \varnothing$. By our construction, there exists some cube Q_t^k such that $10Q_s^{k+1} \subset \frac{3}{2}Q_t^k$, so that Q_s^{k+1} comes from Ω_t^k. Because of the property (3) of the covering, we have $\ell(Q_t^k) \leq 10\ell(Q_j^k)$. Therefore, $2Q_t^k \in \mathcal{AD}_{+\infty, h_j^k-3}$, which implies $Q_s^{k+1} \in \mathcal{AD}_{+\infty, h_j^k+7}$ and $2Q_s^{k+1} \in \mathcal{AD}_{+\infty, h_j^k+6}$.

Let $P \in \mathcal{AD}_{h_j^k+4}$ be some μ-doubling cube whose center is in $\frac{3}{2}Q_j^k$ (which implies $P \subset 2Q_j^k$). Let S_P be the set of indices s such that $2Q_s^{k+1} \cap P \neq \varnothing$. We have $\ell(Q_s^{k+1}) \ll \ell(P)$ for $s \in S_P$, because $2Q_s^{k+1} \in \mathcal{AD}_{+\infty, h_j^k+6}$ (since $2Q_s^{k+1} \cap 2Q_j^k \neq \varnothing$). Thus, $2Q_s^{k+1} \subset 2P$. From our construction, we deduce

$$\mu(\widetilde{A}_{k+2} \cap P) \leq \mu\left(\bigcup_{s \in S_P} (\Omega_s^{k+1} \cap 2Q_s^{k+1})\right) \leq \sum_{s \in S_P} \mu(\Omega_s^{k+1} \cap 2Q_s^{k+1}).$$

Since $N^{h_s^{k+1}+20}\sigma(x) = N^{h_s^{k+1}+20}(\sigma\chi_{3Q_s^{k+1}})(x)$ for $x \in Q_s^{k+1}$, by the weak $(1,1)$ boundedness of $N^{h_s^{k+1}+20}$, and by the σ-doubling property of Q_s^{k+1}, we obtain

$$\mu(\Omega_s^{k+1} \cap 2Q_s^{k+1}) \leq \frac{C\,\sigma(3Q_s^{k+1})}{K\,\lambda_s^{k+1}} \leq \frac{C}{K}\mu(Q_s^{k+1}).$$

Thus, by the finite overlap of the cubes Q_s^{k+1} and the fact that P is μ-doubling,

$$\mu(\widetilde{A}_{k+2} \cap P) \leq \frac{C}{K} \sum_{s \in S_P} \mu(Q_s^{k+1}) \leq \frac{C}{K}\mu(2P) \leq \frac{C}{K}\mu(P) =: \delta_0 \mu(P).$$

Since we may choose K as big as we want, δ_0 can be taken arbitrarily small, and (9.19) holds.

Let us see that (9.18) follows from (9.19). We denote $\widetilde{A}_{k,0}=\bigcup_{j\in I_k^0}2Q_j^k$. Since the cubes $2Q_j^k$, $j\in I_k^0$, are disjoint, (9.19) implies $\sigma(\widetilde{A}_{k,0}\cap\widetilde{A}_{k+2})\leq \tau_1\,\sigma(\widetilde{A}_{k,0})$. By the property (1) of the covering and the fact that Q_j^k is $(100,\beta)$-σ-doubling, we have

$$\sigma(\widetilde{A}_{k,0})=\sum_{j\in I_k^0}\sigma(2Q_j^k)\geq C_{24}^{-1}\sum_{j\in I_k^0}\sigma(40Q_j^k)\geq C_{24}^{-1}\sigma(\widetilde{A}_k).$$

Then,

$$\sigma(\widetilde{A}_k\setminus\widetilde{A}_{k+2})\geq\sigma(\widetilde{A}_{k,0}\setminus\widetilde{A}_{k+2})\geq(1-\tau_1)\,\sigma(\widetilde{A}_{k,0})\geq(1-\tau_1)\,C_{24}^{-1}\sigma(\widetilde{A}_k).$$

Therefore,

$$\sigma(\widetilde{A}_k\cap\widetilde{A}_{k+2})\leq(1-(1-\tau_1)\,C_{24}^{-1})\,\sigma(\widetilde{A}_k)=:\tau_0\sigma(\widetilde{A}_k).$$

The end of the proof. We only need to prove the lemma for $S<+\infty$. For each $k\geq 1$ we have

$$\begin{aligned}\int_{A_k}|N^{H_x^k+20}\sigma(x)|^p\,w(x)\,d\mu(x)&\leq\sum_{i\in I_k}\int_{Q_i^k}|N^{h_i^k+20}\sigma(x)|^p\,w(x)\,d\mu(x)\\&\leq CS\sum_i\sigma(Q_i^k)\leq CS\,\sigma(A_k).\end{aligned}$$

From (9.17) we deduce that $\sigma(A_k)\to 0$ as $k\to\infty$, and then the integral on the left hand side above tends to 0 as $k\to\infty$. Now the lemma follows from (9.16) and (9.17). $\square$

Proof of Lemma 9.1: Let Q be some cube with $Q\in\mathcal{AD}_h$ and $x_0\in Q\cap\operatorname{supp}(\mu)$. We write

$$\int|N(\sigma\chi_Q)|^p\,w\,d\mu=\int_{\frac{21}{20}Q}+\int_{Q_{x_0,h-4}\setminus\frac{21}{20}Q}+\int_{\mathbb{R}^d\setminus Q_{x_0,h-4}}=:I+II+III.$$

First we will estimate the integral I. For each $x\in\frac{21}{20}Q\cap\operatorname{supp}(\mu)$, let P_x be some μ-σ-$(4,\beta)$-doubling cube with $P_x\in\mathcal{AD}_{h+10}$. Notice that for each $y\in P_x$ and $k\geq h+15$, we have $\operatorname{supp}(s_k(y,\cdot))\subset 2P_x$. Thus by Lemma 9.2, if we denote $C_S:=C_\varepsilon+\varepsilon S$, we get

$$\begin{aligned}\int_{P_x}|N^{h+15}\sigma|^p\,w\,d\mu&=\int_{P_x}|N^{h+15}(\sigma\chi_{2P_x})|^p\,w\,d\mu\\&\leq C_S\sigma(2P_x)\leq C\,C_S\,\sigma(P_x).\end{aligned}$$

By Besicovitch's Covering Theorem, there exists some subfamily of cubes $\{P_{x_i}\}_i\subset\{P_x\}_x$ which covers $\frac{21}{20}Q\cap\operatorname{supp}(\mu)$ with finite overlap. Since

$\ell(P_{x_i}) \ll \ell(Q)$, we have $P_{x_i} \subset \frac{11}{10}Q$. Then we obtain

$$\begin{aligned}\int_{\frac{21}{20}Q} |N^{h+15}\sigma|^p \, w \, d\mu &\leq \sum_i \int_{P_{x_i}} |N^{h+15}\sigma|^p \, w \, d\mu \\ &\leq C \, C_S \sum_i \sigma(P_{x_i}) \leq C \, C_S \, \sigma(\tfrac{11}{10}Q).\end{aligned}$$

It is easily seen that, for all $y \in \frac{21}{20}Q$, $N(\sigma\chi_Q)(y) \leq C \, N^{h-2}(\sigma\chi_Q)(y)$. Therefore,

$$\begin{aligned}I &\leq C \int_{\frac{21}{20}Q} |N^{h-2}(\sigma\chi_Q)|^p \, w \, d\mu \\ &\leq C \int_{\frac{21}{20}Q} \left| \sum_{k=h-2}^{h+14} S_k(\chi_Q\sigma) \right|^p w \, d\mu + \int_{\frac{21}{20}Q} |N^{h+15}\sigma|^p \, w \, d\mu \\ &\leq C(1+C_S)\sigma(\tfrac{11}{10}Q).\end{aligned}$$

Now we turn our attention to the integral II. For $y \in Q_{x_0,h-4} \setminus \frac{21}{20}Q$,

$$N(\sigma\chi_Q)(y) \leq \frac{C \, \sigma(Q)}{|y-x_0|^n} \leq C \sum_{k=h-6}^{h+3} S_k(\sigma\chi_Q)(y).$$

Thus,

$$II \leq \int \left| \sum_{k=h-6}^{h+3} S_k(\sigma\chi_Q) \right|^p w \, d\mu \leq C \, \sigma(Q).$$

Finally we deal with III. For $k \leq h-4$ and $y \in Q_{x_0,k-1} \setminus Q_{x_0,k}$, we have

$$N(\sigma\chi_Q)(y) \leq C \frac{\sigma(Q)}{|y-x_0|^n} \leq C \frac{\sigma(Q)}{\sigma(Q_{x_0,k+1})} \sum_{j=k-3}^{k+2} S_j(\sigma\chi_{Q_{x_0,k+1}})(y).$$

Thus,

$$\begin{aligned}\int_{Q_{x_0,k}\setminus Q_{x_0,k-1}} |N(\sigma\chi_Q)|^p \, w \, d\mu &\leq \frac{C\sigma(Q)^p}{\sigma(Q_{x_0,k+1})^p} \int \left| \sum_{j=k-3}^{k+2} S_j(\sigma\chi_{Q_{x_0,k+1}}) \right|^p w \, d\mu \\ &\leq \frac{C\sigma(Q)^p}{\sigma(Q_{x_0,k+1})^{p-1}}.\end{aligned}$$

From Lemma 8.3, we deduce $\sigma(Q) \leq \sigma(Q_{x_0,h-1}) \leq \eta^{h-k-2}\sigma(Q_{x_0,k+1})$ for some fixed constant η with $0 < \eta < 1$. Therefore,

$$\begin{aligned} III &= \sum_{k=-\infty}^{h-4} \int_{Q_{x_0,k-1}\setminus Q_{x_0,k}} |N(\sigma\chi_Q)|^p \, w \, d\mu \\ &\leq C\,\sigma(Q) \sum_{k=-\infty}^{h-4} \eta^{(p-1)(h-k-2)} \leq C\,\sigma(Q). \end{aligned}$$

So we have

$$\int |N(\sigma\chi_Q)|^p \, w \, d\mu \leq C(1 + C_S)\,\sigma(\tfrac{11}{10}Q) = C_{25}(1 + C_\varepsilon + \varepsilon S)\,\sigma(\tfrac{11}{10}Q).$$

Choosing $\varepsilon \leq 1/(2C_{25})$ and taking the supremum over all the cubes Q, we get $S \leq C_{25}(1 + C_\varepsilon) + \frac{1}{2}S$. Thus $S \leq 2C_{25}(1 + C_\varepsilon)$ if $S < +\infty$.

One way to avoid the assumption $S < +\infty$ would be to work with "truncated" operators of type $N^{h,l}f := \sup_{h\leq k\leq l} S_k|f|$ in Lemma 9.2, instead of N^h; and also to consider a truncated version of S in (9.1), etc. The technical details are left for the reader. □

10. Boundedness of N on $L^p(w)$

The implication (e) ⇒ (c) of Lemma 4.2 follows from Lemma 9.1 and the following result.

Lemma 10.1. *If for any $k \in \mathbb{Z}$ and any cube Q,*

$$\int N(\sigma\,\chi_Q)^p \, w \, d\mu \leq C\,\sigma(\tfrac{11}{10}Q) \tag{10.1}$$

and

$$\int S_k(w\,\chi_Q)^{p'} \, \sigma \, d\mu \leq C\, w(Q), \tag{10.2}$$

with C independent of k and Q, then N is bounded on $L^p(w)$.

The proof of this lemma is inspired by the techniques used by Sawyer [**Saw2**] to obtain two weight norm inequalities for fractional integrals. In our case, we have to overcome new difficulties which are mainly due to the fact that the operator N is not linear and it is very far from behaving as a self adjoint operator, because it is a *centered* maximal operator.

Proof: We will show that for some $\beta \geq 0$ the operator $T := N + \beta\, M_R$ is bounded on $L^p(w)$. Without loss of generality, we take $f \in L^1(\mu)$ non negative with compact support. Given some constant $\alpha > 1$ close to 1, for each $k \in \mathbb{Z}$, we denote

$$\Omega_k = \{x : Tf(x) > \alpha^k\}.$$

The precise value of α and β will be fixed below. As in Lemma 7.1, we consider the Whitney decomposition $\Omega_k = \bigcup_i Q_i^k$, where Q_i^k are dyadic cubes with disjoint interiors (the *Whitney cubes*).

Take some cube $Q_i^k \subset \Omega_k$ and $x \in Q_i^k \cap \Omega_{k+2}$. Suppose that m and β are chosen in Lemma 7.1 so that the maximum principle (7.1) holds with $\varepsilon = \alpha - 1$. Then, we have

$$T(f\,\chi_{\mathbb{R}^d \setminus U_m(Q_i^k)})(x) \leq (1+\varepsilon)\,\alpha^k = \alpha^{k+1}, \tag{10.3}$$

and so

$$T(f\,\chi_{U_m(Q_i^k)})(x) \geq \alpha^{k+2} - \alpha^{k+1} = \frac{\alpha-1}{\alpha}\alpha^{k+2}. \tag{10.4}$$

Let $h \in \mathbb{Z}$ be such that $Q_i^k \in \mathcal{AD}_h$. If for all j with $h - M \leq j \leq h + M$ (where M is some positive big integer which will be chosen later) we have

$$S_j(f\chi_{U_m(Q_i^k)})(x) \leq \delta\,\alpha^k,$$

where $\delta > 0$ is another constant which we will choose below, then we write $x \in B_k$ (i.e. x is a "bad point"). Notice that $B_k \subset \Omega_{k+2} \subset \Omega_k$.

We will see that the set of bad points is quite small. Indeed, we will prove that

$$w\left(\bigcup_{j \geq k} B_j\right) \leq \eta\, w(\Omega_k), \tag{10.5}$$

where $0 < \eta < 1$ is some constant which depends on τ (from the Z_∞ property), n, d, but not on β, m, α, M. We defer the proof of (10.5), which is one of the key points of our argument, until Lemma 10.2 below.

Let us denote $A_k = \bigcup_{j\geq k} B_j$. Now we have

$$\begin{aligned}
\int |Tf|^p\, w\, d\mu &= \int_0^\infty p\,\lambda^{p-1}\, w(\Omega_\lambda)\, d\lambda \\
&\leq \sum_{k\in\mathbb{Z}} p\,(\alpha^{k+1} - \alpha^k)\,\alpha^{(k+1)(p-1)}\, w(\Omega_k) \qquad (10.6)\\
&= p\,\alpha^{p-1}(\alpha-1)\sum_{k\in\mathbb{Z}} \alpha^{kp}\big[w(\Omega_k\setminus A_{k-2}) + w(\Omega_k \cap A_{k-2})\big].
\end{aligned}$$

From (10.5) we get

$$p\,\alpha^{p-1}(\alpha-1)\sum_{k\in\mathbb{Z}}\alpha^{kp}w(\Omega_k\cap A_{k-2}) \leq \eta\,p\,\alpha^{p-1}(\alpha-1)\sum_{k\in\mathbb{Z}}\alpha^{kp}\,w(\Omega_{k-2}). \tag{10.7}$$

From calculations similar to the ones in (10.6), it follows

$$\int |Tf|^p\,w\,d\mu \geq p\,(\alpha-1)\,\alpha^{-3p}\sum_{k\in\mathbb{Z}}\alpha^{kp}\,w(\Omega_{k-2}).$$

If we take α such that $\eta^{1/2}\alpha^{4p-1}=1$, then the right hand side of (10.7) is bounded above by $\eta^{1/2}\int|Tf|^p\,w\,d\mu$, and so

$$\int |Tf|^p\,w\,d\mu \leq (1-\eta^{1/2})^{-1}\sum_{k\in\mathbb{Z}}\alpha^{kp}\,w(\Omega_k\setminus A_{k-2}).$$

Summing by parts we get

$$\begin{aligned}\int |Tf|^p\,w\,d\mu &\leq C\sum_{k\in\mathbb{Z}}\alpha^{kp}\,w(\Omega_{k+2}\setminus A_k)\\ &= C\sum_{k\in\mathbb{Z}}\alpha^{kp}\,\big[w(\Omega_{k+2}\setminus A_k)-w(\Omega_{k+3}\setminus A_{k+1})\big].\end{aligned}$$

Observe that if we assume $\int|Tf|^p\,w\,d\mu<\infty$, then

$$\sum_{k\in\mathbb{Z}}\alpha^{kp}\,w(\Omega_{k+2}\setminus A_k)<\infty \quad\text{and}\quad \sum_{k\in\mathbb{Z}}\alpha^{kp}w(\Omega_{k+3}\setminus A_{k+1})<\infty,$$

which implies that our summation by parts is right. Since $A_{k+1}\subset A_k$, we have $w(\Omega_{k+3}\setminus A_{k+1})\geq w(\Omega_{k+3}\setminus A_k)$. Thus,

$$\int |Tf|^p\,w\,d\mu \leq C\sum_{k\in\mathbb{Z}}\alpha^{kp}\,w((\Omega_{k+2}\setminus\Omega_{k+3})\setminus A_k). \tag{10.8}$$

We denote $E_i^k=Q_i^k\cap(\Omega_{k+2}\setminus\Omega_{k+3})\setminus A_k$ for all (k,i). To simplify notation, we also set $U_i^k=U_m(Q_i^k)$. Given $h\in\mathbb{Z}$ such that $Q_i^k\in\mathcal{AD}_h$, we consider the operator

$$\mathcal{S}_{(k,i)}=S_{h-M}+S_{h-M+1}+\cdots+S_{h+M}.$$

Since $E_i^k \subset \mathbb{R}^d \setminus A_k$, we obtain

$$\begin{aligned} w(E_i^k) &\le \delta^{-1}\alpha^{-k} \int_{E_i^k} \mathcal{S}_{(k,i)}(\chi_{U_i^k} f)\, w\, d\mu \\ &= \delta^{-1}\alpha^{-k} \int_{U_i^k} f\, \mathcal{S}^*_{(k,i)}(\chi_{E_i^k} w)\, d\mu \\ &= \delta^{-1}\alpha^{-k} \left(\int_{U_i^k \setminus \Omega_{k+3}} + \int_{U_i^k \cap \Omega_{k+3}} \right) = \delta^{-1}\alpha^{-k}(\sigma_i^k + \tau_i^k). \end{aligned}$$

From (10.8) we get

$$\begin{aligned} \int |Tf|^p\, w\, d\mu &\le C \sum_{k,i} \alpha^{kp}\, w(E_i^k) \\ &= C \left(\sum_{(k,i)\in E} + \sum_{(k,i)\in F} + \sum_{(k,i)\in G} \right) \cdot \alpha^{kp}\, w(E_i^k) \qquad (10.9) \\ &= C\,(I + II + III), \end{aligned}$$

where

$$\begin{aligned} E &= \left\{(k,i) : w(E_i^k) \le \theta\, w(Q_i^k)\right\}, \\ F &= \left\{(k,i) : w(E_i^k) > \theta\, w(Q_i^k) \text{ and } \sigma_i^k > \tau_i^k\right\}, \\ G &= \left\{(k,i) : w(E_i^k) > \theta\, w(Q_i^k) \text{ and } \sigma_i^k \le \tau_i^k\right\}, \end{aligned}$$

and where θ is some constant with $0 < \theta < 1$ which will be chosen below.

The term I is easy to estimate:

$$\begin{aligned} I &= \sum_{(k,i)\in E} \alpha^{kp}\, w(E_i^k) \le \theta \sum_{k,i} \alpha^{kp}\, w(Q_i^k) \\ &\le \theta \sum_k \alpha^{kp}\, w\{Tf > \alpha^k\} \le C\,\theta \int |Tf|^p\, w\, d\mu. \end{aligned}$$

Let us consider the term II now. By Hölder's inequality and (10.2), we obtain

$$\begin{aligned} II &= \sum_{(k,i)\in F} \alpha^{kp}\, w(E_i^k) \leq \sum_{(k,i)\in F} w(E_i^k) \left(\frac{2\sigma_i^k}{\delta w(E_i^k)} \right)^p \\ &\leq C\theta^{-p}\delta^{-p} \sum_{k,i} w(E_i^k) \left(\frac{1}{w(Q_i^k)} \int_{U_i^k\setminus\Omega_{k+3}} f\, \mathcal{S}^*_{(k,i)}(w\,\chi_{E_i^k})\, d\mu \right)^p \\ &\leq C\theta^{-p}\delta^{-p} \sum_{k,i} \frac{w(E_i^k)}{w(Q_i^k)^p} \left(\int_{U_i^k} \mathcal{S}^*_{(k,i)}(w\,\chi_{E_i^k})^{p'}\,\sigma\, d\mu \right)^{p/p'} \left(\int_{U_i^k\setminus\Omega_{k+3}} f^p w\, d\mu \right) \\ &\leq C\theta^{-p}\delta^{-p} \sum_{k,i} \int_{U_i^k\setminus\Omega_{k+3}} f^p\, w\, d\mu. \end{aligned}$$

It is easy to check that the family of sets $\{U_i^k\}_i$ has finite overlap for each k, with some constant which depends on m. This fact implies

$$\sum_{k,i} \chi_{U_i^k\setminus\Omega_{k+3}} \leq C \sum_k \chi_{\Omega_k\setminus\Omega_{k+3}} \leq C.$$

Therefore, $II \leq C \int f^p\, w\, d\mu$.

Finally we have to deal with the term III. Notice that $E_i^k \subset \mathbb{R}^d\setminus\Omega_{k+3}$ and, for $y \notin \Omega_{k+3}$, by Lemmas 8.1 and 3.7 we have

$$\sup_{x\in 2Q_j^{k+3}} s_t(y,x) \leq C \inf_{x\in 2Q_j^{k+3}} \sum_{r=t-6}^{t+6} s_r(y,x) \tag{10.10}$$

for all $t \in \mathbb{Z}$. Let $H_i^k = \{j : Q_j^{k+3}\cap U_i^k \neq \varnothing\}$. Then, for $j \in H_i^k$, we have

$$\sup_{x\in 2Q_j^{k+3}} \mathcal{S}^*_{(k,i)}(w\,\chi_{E_i^k})(x) \leq C \inf_{x\in 2Q_j^{k+3}} \sum_{t=h-M-6}^{h+M+6} S_t^*(w\,\chi_{E_i^k})(x).$$

We set $\overline{\mathcal{S}}_{(k,i)} = \sum_{t=h-M-6}^{h+M+6} S_t$, and we obtain

$$\begin{aligned}\tau_i^k &= \int_{U_i^k \cap \Omega_{k+3}} f\, \mathcal{S}_{(k,i)}(w\,\chi_{E_i^k})\, d\mu \\ &\le C \sum_{j\in H_i^k} \inf_{x\in 2Q_j^{k+3}} \overline{\mathcal{S}}^*_{(k,i)}(w\,\chi_{E_i^k})(x) \int_{Q_j^{k+3}} f\, d\mu \\ &\le C \sum_{j\in H_i^k} \left(\int_{2Q_j^{k+3}} \overline{\mathcal{S}}^*_{(k,i)}(w\,\chi_{E_i^k})\,\sigma\, d\mu\right)\left(\frac{1}{\sigma(2Q_j^{k+3})}\int_{Q_j^{k+3}} f\, d\mu\right).\end{aligned}$$

We denote $T_i^k = \int_{Q_i^k} f\, d\mu/\sigma(2Q_i^k)$ and $L_i^k = \{s : Q_s^k \cap U_i^k \neq \varnothing\}$. Then we have

$$\begin{aligned}\tau_i^k &\le C \sum_{j\in H_i^k} \left(\int_{2Q_j^{k+3}} \overline{\mathcal{S}}^*_{(k,i)}(w\,\chi_{E_i^k})\,\sigma\, d\mu\right) T_j^{k+3} \\ &\le C \sum_{s\in L_i^k} \sum_{j:Q_j^{k+3}\subset Q_s^k} \left(\int_{2Q_j^{k+3}} \overline{\mathcal{S}}^*_{(k,i)}(w\,\chi_{E_i^k})\,\sigma\, d\mu\right) T_j^{k+3}.\end{aligned} \tag{10.11}$$

We will show that

$$\sum_{\substack{(k,i)\in G\\ k\ge N_0\\ k=M_0 \mod 3}} \alpha^{kp}\, w(E_i^k) \le C \int f^p w\, d\mu, \tag{10.12}$$

for any N_0 and M_0. For the rest of the proof we follow the convention that all indices (k,i) are restricted to $k \ge N_0$ and $k = M_0 \mod 3$.

Now we will introduce *principal* cubes as in [**Saw2**, p. 540] or [**MW**, p. 804]. Let G_0 be the set of indices (k,i) such that Q_i^k is maximal. Assuming G_n already defined, G_{n+1} consists of those (k,i) for which there is $(t,u) \in G_n$ with $Q_i^k \subset Q_u^t$ and

(a) $T_i^k > 2T_u^t$,
(b) $T_s^r \le 2T_u^t$ if $Q_i^k \subsetneq Q_s^r \subset Q_u^t$.

We denote $\Gamma = \bigcup_{n=0}^\infty G_n$, and for each (k,i), we define $P(Q_i^k)$ as the smallest cube Q_u^t containing Q_i^k with $(t,u)\in\Gamma$. Then we have

(a) $P(Q_i^k) = Q_u^t$ implies $T_i^k \le 2T_u^t$.
(b) $Q_i^k \subsetneq Q_u^t$ and (k,i), $(t,u) \in \Gamma$ implies $T_i^k > 2T_u^t$.

By (10.11) and the fact that $\#L_i^k \leq C$, we get

$$
\begin{aligned}
&\sum_{(k,i)\in G} \alpha^{kp}\, w(E_i^k) \\
&\leq \sum_{(k,i)\in G} w(E_i^k) \left(\frac{2\tau_i^k}{\delta w(E_i^k)}\right)^p \leq C \sum_{k,i} \frac{w(E_i^k)}{w(Q_i^k)^p}\,(\tau_i^k)^p \\
&\leq C \sum_{k,i} \sum_{s\in L_i^k} \frac{w(E_i^k)}{w(Q_i^k)^p} \Bigg[\sum_{\substack{j:(k+3,j)\notin\Gamma \\ Q_j^{k+3}\subset Q_s^k \\ P(Q_j^{k+3})=P(Q_s^k)}} \left(\int_{2Q_j^{k+3}} \overline{\mathcal{S}}^*_{(k,i)}(w\,\chi_{E_i^k})\,\sigma\, d\mu\right) T_j^{k+3}\Bigg]^p \\
&\quad + C \sum_{k,i} \frac{w(E_i^k)}{w(Q_i^k)^p} \Bigg[\sum_{j\in H_i^k:(k+3,j)\in\Gamma} \left(\int_{2Q_j^{k+3}} \overline{\mathcal{S}}^*_{(k,i)}(w\,\chi_{E_i^k})\,\sigma\, d\mu\right) T_j^{k+3}\Bigg]^p \\
&= IV + V.
\end{aligned}
$$

Let us estimate the term IV first. Notice that if $(k+3,j)\notin\Gamma$, then $Q_j^{k+3} \neq P(Q_j^{k+3})$. As a consequence, $\ell(Q_j^{k+3}) \leq \ell(P(Q_j^{k+3}))/2$, and $2Q_j^{k+3} \subset \frac{4}{3}P(Q_j^{k+3})$. Taking into account this fact, the finite overlap of the cubes Q_j^{k+3} (for a fixed k), and (10.1), for any $(t,u)\in\Gamma$ we get

$$
\begin{aligned}
&\sum_{k,i} \sum_{s\in L_i^k:P(Q_s^k)=Q_u^t} \frac{w(E_i^k)}{w(Q_i^k)^p} \Bigg[\sum_{\substack{j:(k+3,j)\notin\Gamma \\ Q_j^{k+3}\subset Q_s^k \\ P(Q_j^{k+3})=Q_u^t}} \left(\int_{2Q_j^{k+3}} \overline{\mathcal{S}}^*_{(k,i)}(w\,\chi_{E_i^k})\,\sigma\, d\mu\right) T_j^{k+3}\Bigg]^p \\
&\leq C \sum_{k,i} \sum_{s\in L_i^k:P(Q_s^k)=Q_u^t} w(E_i^k) \left(\frac{1}{w(Q_i^k)} \int_{\frac{4}{3}Q_u^t} \overline{\mathcal{S}}^*_{(k,i)}(w\,\chi_{Q_i^k})\,\sigma\, d\mu\right)^p (2T_u^t)^p \\
&\leq C(T_u^t)^p \sum_{k,i} \sum_{s\in L_i^k:P(Q_s^k)=Q_u^t} w(E_i^k) \left(\frac{1}{w(Q_i^k)} \int_{Q_i^k} \overline{\mathcal{S}}_{(k,i)}(\sigma\,\chi_{\frac{4}{3}Q_u^t})\, w\, d\mu\right)^p \\
&\leq C(T_u^t)^p \int M_w^d(N(\sigma\,\chi_{Q_u^t}))^p\, w\, d\mu \\
&\leq C(T_u^t)^p \int N(\sigma\,\chi_{\frac{4}{3}Q_u^t})^p\, w\, d\mu \leq C(T_u^t)^p\, \sigma(2Q_u^t),
\end{aligned}
$$

where we have denoted by M_w^d the dyadic maximal operator with respect to w. Thus,

$$IV \le C \sum_{(t,u)\in\Gamma} \sigma(2Q_u^t)(T_u^t)^p. \tag{10.13}$$

Let us estimate the term V. By Hölder's inequality and (10.2), for a fixed (k,i),

$$\begin{aligned}
&\frac{w(E_i^k)}{w(Q_i^k)^p}\left[\sum_{j\in H_i^k:(k+3,j)\in\Gamma}\left(\int_{2Q_j^{k+3}}\overline{\mathcal{S}}^*_{(k,i)}(w\,\chi_{E_i^k})\,\sigma\,d\mu\right)\,T_j^{k+3}\right]^p\\
&\qquad\le\frac{w(E_i^k)}{w(Q_i^k)^p}\left[\sum_{j\in H_i^k}\sigma(2Q_j^{k+3})^{-p'/p}\left(\int_{2Q_j^{k+3}}\overline{\mathcal{S}}^*_{(k,i)}(w\,\chi_{E_i^k})\,\sigma\,d\mu\right)^{p'}\right]^{p/p'}\\
&\qquad\quad\times\left[\sum_{j\in H_i^k:(k+3,j)\in\Gamma}\sigma(2Q_j^{k+3})\,(T_j^{k+3})^p\right]\\
&\qquad\le\frac{w(E_i^k)}{w(Q_i^k)^p}\left[\sum_{j\in H_i^k}\int_{2Q_j^{k+3}}\overline{\mathcal{S}}^*_{(k,i)}(w\,\chi_{E_i^k})^{p'}\,\sigma\,d\mu\right]^{p/p'}\\
&\qquad\quad\times\left[\sum_{j\in H_i^k:(k+3,j)\in\Gamma}\sigma(2Q_j^{k+3})\,(T_j^{k+3})^p\right]\\
&\qquad\le C\sum_{j\in H_i^k:(k+3,j)\in\Gamma}\sigma(2Q_j^{k+3})(T_j^{k+3})^p.
\end{aligned}$$

Summing over (k,i), since any cube Q_j^{k+3} occurs at most C times in the resulting sum, we get

$$V \le C \sum_{(t,u)\in\Gamma} \sigma(2Q_u^t)(T_u^t)^p. \tag{10.14}$$

Notice that for each (t,u) we can write

$$\begin{aligned}
\sigma(2Q_u^t)(T_u^t)^p &= \sigma(Q_u^t)(T_u^t)^{p-1}\frac{1}{\sigma(Q_u^t)}\int_{Q_u^t} f\sigma^{-1}\,\sigma\,d\mu\\
&=: \sigma(Q_u^t)(T_u^t)^{p-1}m_{\sigma,Q_u^t}(f\sigma^{-1}).
\end{aligned}$$

We have obtained

$$IV + V \le C \sum_{(t,u)\in\Gamma} \sigma(Q_u^t)(T_u^t)^{p-1} m_{\sigma,Q_u^t}(f\sigma^{-1})$$
$$= C \int \Biggl(\sum_{(t,u)\in\Gamma} (T_u^t)^{p-1} m_{\sigma,Q_u^t}(f\sigma^{-1})\, \chi_{Q_u^t}(x) \Biggr) \sigma(x)\, d\mu(x).$$

Notice that for any fixed x we have

$$\sum_{(t,u)\in\Gamma} (T_u^t)^{p-1} m_{\sigma,Q_u^t}(f\sigma^{-1}) \chi_{Q_u^t}(x) \le C \sup_{(t,u)\in\Gamma: x\in Q_u^t} (T_u^t)^{p-1} M_\sigma^d(f\sigma^{-1})(x)$$
$$\le C\, M_\sigma^d(f\sigma^{-1})(x)^p.$$

Therefore,

$$IV + V \le C \int M_\sigma^d(f\sigma^{-1})^p\, \sigma\, d\mu \le C \int (f\sigma^{-1})^p\, \sigma\, d\mu = C \int f^p\, w\, d\mu,$$

which yields (10.12). Thus, by (10.9),

$$\int |Tf|^p\, w\, d\mu \le C(I + II + III) \le C\theta \int |Tf|^p\, w\, d\mu + C \int f^p\, w\, d\mu.$$

We only have to choose θ small enough, and we are done. $\square$

To complete the proof of the implication (d) $\Rightarrow$ (c) of Lemma 4.2, it remains to prove the following result.

Lemma 10.2. *With the notation and assumptions of Lemma 10.1, (10.5) holds. That is,* $w\Bigl(\bigcup_{j\ge k} B_j\Bigr) \le \eta\, w(\Omega_k)$, *with* $0 < \eta < 1$.

Before proving the lemma, a remark:

Remark 10.3. Besicovitch's Covering Theorem asserts that if $A \subset \mathbb{R}^d$ is bounded and there exists some family of cubes $\mathcal{Q} = \{Q_x\}_{x\in A}$, with each Q_x centered at x, then there exists some finite or countable family of cubes $\{Q_{x_i}\}_i \subset \mathcal{Q}$ which covers A with finite overlap. That is, $\chi_A \le \sum_i \chi_{Q_{x_i}} \le C$, with C depending only on d.

We are going to show that the covering $\{Q_{x_i}\}_i$ can be chosen so that the following property holds too:

(10.15) *If* $z \in A \cap Q_{x_i}$ *for some* i, *then* $\ell(Q_z) \le 4\ell(Q_{x_i})$.

Indeed, for each $x \in A$, let R_x be some cube of the type Q_y, $y \in A$, with $x \in \frac{1}{2}R_x$ and such that

$$\ell(R_x) \geq \frac{99}{100} \sup_{y: x \in \frac{1}{2}Q_y} \ell(Q_y).$$

Now we will apply Besicovitch's Covering Theorem to the family of cubes $\{R_x\}_{x\in A}$. Let us remark that the Theorem of Besicovitch also holds for the family $\{R_x\}_{x\in A}$ because, although the cubes R_x are not centered at x, we still have $x \in \frac{1}{2}R_x$ (see [**Mo**] or [**Gu**, pp. 6–7], for example). So there exists some finite or countable family $\{R_{x_i}\}_i$ which covers A with finite overlap. Notice that $\{R_{x_i}\}_i \subset \mathcal{Q}$, and if $z \in R_{x_i} \cap A$, then $\ell(Q_z) \leq 4\ell(R_{x_i})$. Otherwise, $x_i \in \frac{1}{2}Q_z$ and $\ell(Q_z) > 4\ell(R_{x_i})$, which contradicts the definition of R_{x_i}.

It is worth comparing this version of Besicovitch Covering Theorem with the version of Wiener's Covering Lemma 8.4. Notice that the statement (3) of Lemma 8.4 and (10.15) look quite similar.

Proof of Lemma 10.2: We use the same notation as in the proof of the preceding lemma.

Let $x \in B_j$ and take Q_i^j containing x (recall $B_j \subset \Omega_{j+2} \subset \Omega_j$), with $Q_i^j \in \mathcal{AD}_h$. By (10.4), we have $N(f\,\chi_{U_i^j})(x) \geq \varepsilon_0\,\alpha^j$ for some $\varepsilon_0 > 0$ depending on α, β, m. It is easily seen that this implies that $S_t(f\,\chi_{U_i^j})(x) \geq \varepsilon_0\,\alpha^j$ for $t \geq h - M$, where M is some positive constant which depends on $\ell(U_i^j)/\ell(Q_i^j)$. Also, by the definition of B_j, $S_t(f\,\chi_{U_i^j})(x) \leq \delta\alpha^j$ for $h - M \leq t \leq h + M$. So, if we choose $\delta \leq \varepsilon_0$ and $M \geq 10$, then

$$\sup_{t \geq h+10} S_t f(x) > \varepsilon_0\,\alpha^j. \tag{10.16}$$

We denote $A_k := \bigcup_{j\geq k} B_j$. For a fixed $x \in A_k$, let r be the least integer such that $r \geq k$ and $x \in B_r$. There exists some cube Q_i^r containing x, with $Q_i^r \in \mathcal{AD}_h$ for some h. Since $S_{h+5}(f\,\chi_{U_i^r})(x) \leq \delta\alpha^r$, by Lemma 3.11 there exists some doubling cube $P_x \in \mathcal{AD}_{h+5,h+4}$ centered at x such that

$$\frac{1}{\mu(2P_x)} \int_{2P_x} f\,d\mu \leq C\delta\alpha^r. \tag{10.17}$$

Now, by Besicovitch's Covering Theorem, we can find some family of cubes $\{P_{x_s}\}_s$ (with $x_s \in A_k$) which covers A_k with finite overlap. Moreover, we assume that the covering has been chosen so that the property (10.15) holds.

Given any ρ with $0 < \rho < 1$, we will show that if δ is small enough, then

$$\mu(A_k \cap P_{x_s}) \leq \rho\, \mu(P_{x_s}) \tag{10.18}$$

for all s.

Let P_{x_0} some fixed cube from the family $\{P_{x_s}\}_s$, and let r_0 be the least integer such that $x_0 \in B_{r_0}$. First we will see that

$$\mu\biggl(\bigcup_{j\geq r_0} B_j \cap P_{x_0}\biggr) \leq \frac{\rho}{2}\mu(P_{x_0}). \tag{10.19}$$

If $z \in B_j \cap P_{x_0}$ for some $j \geq r_0$ and $z \in Q_i^j$, then by (10.16) we have

$$\sup_{j\geq h+10} S_j f(z) > \varepsilon_0\, \alpha^j, \tag{10.20}$$

where h is so that $Q_i^j \in \mathcal{AD}_h$. Let us denote by $Q_{i_0}^{r_0}$ the Whitney cube of Ω_{r_0} containing x_0, with $Q_{i_0}^{r_0} \in \mathcal{AD}_{h_0}$. Since $\Omega_j \subset \Omega_{r_0}$, we have $\ell(Q_i^j) \leq C_{26}\ell(Q_{i_0}^{r_0})$, and so $h \geq h_0 - 1$. In fact, if C_{26}, which depends on d, is very big, then we should write $h \geq h_0 - q$, where q is some positive integer big enough, depending on C_{26}. The details of the required modifications in this case are left to the reader. From (10.20), we get

$$\sup_{j\geq h_0+9} S_j f(z) > \varepsilon_0\, \alpha^j \geq \varepsilon_0\, \alpha^{r_0}. \tag{10.21}$$

For $j \geq h_0 + 9$ and $z \in P_{x_0}$, we have $\operatorname{supp}(s_j(z,\cdot)) \subset 2P_{x_0}$, because $P_{x_0} \in \mathcal{AD}_{h_0+5,h_0+4}$. Thus (10.21) implies $N(f\,\chi_{2P_{x_0}})(z) > \varepsilon_0\, \alpha^{r_0}$, and then, from the weak $(1,1)$ boundedness of N, by (10.17), and because P_{x_0} is doubling, we obtain

$$\begin{aligned}\mu\biggl(\bigcup_{j\geq r_0} B_j \cap P_{x_0}\biggr) &\leq \mu\left\{z \in P_{x_0} : N(f\,\chi_{2P_{x_0}})(z) > \varepsilon_0\, \alpha^{r_0}\right\}\\ &\leq \frac{C}{\varepsilon_0\, \alpha^{r_0}} \int_{2P_{x_0}} f\, d\mu \leq C\, \varepsilon_0^{-1}\, \delta\, \mu(P_{x_0}).\end{aligned} \tag{10.22}$$

So (10.19) holds if δ is sufficiently small.

Now we have to estimate $\mu\Bigl(\bigcup_{k\leq j\leq r_0-1} B_j \cap P_{x_0}\Bigr)$. If $z \in B_j \cap P_{x_0}$, then $\ell(P_z) \leq 4\,\ell(P_{x_0})$, by (10.15). Recall also that $P_{x_0} \in \mathcal{AD}_{h_0+5,h_0+4}$. As a consequence, we deduce $P_z \in \mathcal{AD}_{+\infty,h_0+3}$. Moreover, we have $P_{x_0} \subset \{Tf > \alpha^{r_0}\}$, and so $Nf(z) > C_{27}\alpha^{r_0}$, with $C_{27} > 0$. Since by (10.3) we have

$$N(f\,\chi_{\mathbb{R}^d\setminus U_i^j})(z) \leq C_{28}\, \alpha^j,$$

we obtain

$$N(f\,\chi_{U_i^j})(z) > C_{27}\alpha^{r_0} - C_{28}\alpha^j \geq \frac{C_{27}}{2}\,\alpha^{r_0},$$

assuming $j \leq r_0 - r_1$, where r_1 is some positive integer which depends on C_{27} and C_{28}. Recall also that the fact that $z \in B_j$ yields

$$S_t(f\,\chi_{U_i^j})(z) \leq \delta\,\alpha^j \quad \text{for } h_1 - 10 \leq t \leq h_1 + 10, \tag{10.23}$$

where h_1 is given by $Q_i^j \in \mathcal{AD}_{h_1}$. If we choose δ small enough, then $\delta\,\alpha^j \leq C_{27}\alpha^{r_0}/2$ and, for $j \leq r_0 - r_1$, (10.23) implies

$$S_t(f\,\chi_{U_i^j})(z) > \frac{C_{27}}{2}\alpha^{r_0} \quad \text{for some } t \geq h_1 + 10. \tag{10.24}$$

On the other hand, if $r_0 - r_1 < j < r_0$, then by (10.16) we have

$$S_t f(z) > \varepsilon_0\,\alpha^{r_0 - r_1} \geq C_{29}\alpha^{r_0} \quad \text{for some } t \geq h_1 + 10, \tag{10.25}$$

with $C_{29} > 0$.

In any case, from the fact that $P_z \in \mathcal{AD}_{h_1+5,h_1+4}$ we deduce $h_1 \geq h_0 - 2$, and so $\operatorname{supp}(s_t(z,\cdot)) \subset 2P_{x_0}$ for $t \geq h_1 + 10$. Thus, from (10.24) and (10.25) we get

$$N(f\,\chi_{2P_{x_0}})(z) \geq \min(C_{27}/2,\, C_{29})\,\alpha^{r_0}$$

for any j with $k \leq j < r_0$. If we take δ small enough, operating as in (10.22), we obtain

$$\mu\Biggl(\bigcup_{k\leq j<r_0} B_j \cap P_{x_0}\Biggr) \leq C\delta\mu(P_{x_0}) \leq \frac{\rho}{2}\mu(P_{x_0}),$$

which together with (10.19) implies (10.18).

By (10.18) and Lemma 5.3, using the Z_∞ condition for w, we get $w(2Q_i^k \setminus A_k) \geq \tau\,w(Q_i^k)$ for each Whitney cube $Q_i^k \in \Omega_k$. By the finite overlap of the cubes $2Q_i^k$, we obtain

$$\tau\,w(\Omega_k) \leq \tau \sum_i w(Q_i^k) \leq \sum_i w(2Q_i^k \setminus A_k) \leq C_{30} w(\Omega_k \setminus A_k).$$

Therefore,

$$w(A_k) \leq (1 - C_{30}^{-1}\tau)\,w(\Omega_k) =: \eta\,w(\Omega_k). \qquad \square$$

11. The general case

In this section we consider the case where not all the cubes $Q_{x,k} \in \mathcal{D}$ are transit cubes.

If $\mathbb{R}^d$ is an initial cube but there are no stopping cubes, then the arguments in Sections 5–10 with some minor modifications are still valid.

If there exist stopping cubes, some problems arise because the functions $S_k\chi_{\mathbb{R}^d}$ are not bounded away from zero, in general. As a consequence, the property Z_∞ has to be modified. Indeed, notice that if we set $A := \mathbb{R}^d$ and Q is some cube which contains stopping points, then (5.1) may fail, and so the Z_∞ condition is useless in this case.

The new formulation of the Z_∞ property is the following. For $k \in \mathbb{Z}$, we denote $\mathcal{ST}_k := \{x \in \operatorname{supp}(\mu) : Q_{x,k} \text{ is a stopping cube}\}$. Notice by the way that $S_j f(x) = 0$ for $j \geq k+2$ and $x \in \mathcal{ST}_k$.

Definition 11.1. We say that w satisfies the Z_∞ property if there exists some constant $\tau > 0$ such that for any cube $Q \in \mathcal{AD}_k$ and any set $A \subset \mathbb{R}^d$ with $Q \cap \mathcal{ST}_{k+3} \subset A$, if

$$S_{k+3}\chi_A(x) \geq 1/4 \quad \text{for all } x \in Q \setminus \mathcal{ST}_{k+3}, \tag{11.1}$$

then $w(A \cap 2Q) \geq \tau\, w(Q)$.

With this new definition, Lemma 5.2 still holds. The new proof is a variation of the former one. On the other hand, Lemma 5.3 changes. Let us state the new version:

Lemma 11.2. *Suppose that w satisfies the Z_∞ property. Let $Q \in \mathcal{AD}_h$ and $A \subset \mathbb{R}^d$ be such that $A \cap Q \cap \mathcal{ST}_{h+4} = \varnothing$. Let $\{P_i\}_i$ be a family of cubes with finite overlap such that $A \cap \frac{3}{2}Q \subset \bigcup_i P_i$, with $P_i \in \mathcal{AD}_{+\infty,h+4}$ and $\ell(P_i) > 0$ for all i. There exists some constant $\delta > 0$ such that if $\mu(A \cap P_i) \leq \delta\, \mu(P_i)$ for each i, then*

$$w(2Q \setminus A) \geq \tau\, w(Q), \tag{11.2}$$

for some constant $\tau > 0$ (depending on Z_∞). If, moreover, $w(2Q) \leq C_{11} w(Q)$, then

$$w(A \cap 2Q) \leq (1 - C_{11}^{-1}\tau)\, w(2Q). \tag{11.3}$$

The proof is analogous to the proof of Lemma 5.3, and it is left for the reader.

The results stated in the other lemmas in Sections 5–10 remain true in the new situation. However, the use of the Z_∞ condition is basic in the proofs of Lemma 5.4, the implication (e) $\Rightarrow$ (c) of Lemma 4.1, Lemma 9.2, Lemma 10.1, and Lemma 10.2. Below we will describe the changes required in the arguments. In the rest of the lemmas and results, the proofs and arguments either are identical or require only some minor modifications (which are left for the reader again).

Changes in the proof of Lemma 5.4. The proof is the same until (5.5), which still holds. Given $Q_i \in \mathcal{AD}_k$, it is easily seen that if $y \in \mathcal{ST}_{k+3} \cap Q_i$, then $T_*(f\chi_{3Q_i})(y) \le C_{31} Nf(y)$. By (5.5), if we choose $\delta < \varepsilon/2C_{31}$, then $A_\lambda \cap Q_i \cap \mathcal{ST}_{k+3} = \varnothing$.

On the other hand, now the estimate (5.6) is valid for $y \in Q_i \setminus \mathcal{ST}_{k+3}$. Then we deduce $S_{k+3}\chi_{2Q_i \setminus A_\lambda}(y) > \frac{1}{4}$ for $y \in Q_i \setminus \mathcal{ST}_{k+3}$, and by the Z_∞, condition we get $w(2Q_i \setminus A_\lambda) \ge \tau w(Q_i)$. Arguing as in (5.7), we obtain $w(A_\lambda) \le \rho, w(\Omega_\lambda)$.

Changes in the proof of the implication (e) ⇒ (c) of Lemma 4.1. The sets Ω_λ, G_λ and B_λ are defined in the same way. The estimates for $w(Q_i \cap G_\lambda)$ are the same.

As shown in (7.11), if $z \in B_\lambda \cap Q_i$, with $Q_i \in \mathcal{AD}_h$, then $S_k(f\chi_{U_i})(z) \ge C_{16}\lambda \ne 0$ for some $k \ge h+6$. This implies $z \notin \mathcal{ST}_{h+4}$. Now the arguments used to prove that $w(B_\lambda) \le \eta_1 w(\Omega_\lambda)$ are still valid, because $B_\lambda \cap Q_i \cap \mathcal{ST}_{h+4} = \varnothing$.

Changes in the proof of Lemma 9.2.

The construction. The construction is basically the same. The only difference is that now we must be careful because the cubes Q_x^1 (and Q_x^k for $k > 1$) may fail to exist due to the existence of stopping points. In the first step of the construction ($k = 1$), we circumvent this problem as follows. If $x \in R_j^1 \setminus \mathcal{ST}_{g_j^1+18}$, then we take a μ-σ-$(100, \beta)$-doubling cube $Q_x^1 \in \mathcal{AD}_{g_j^1+16}$. If $x \in R_j^1 \cap \mathcal{ST}_{g_j^1+18}$, we write $x \in \mathcal{AS}_1$. We consider a Besicovitch covering of $Q_0 \cap \Omega_0 \setminus \mathcal{AS}_1$ with this type of cubes: $Q_0 \cap \Omega_0 \setminus \mathcal{AS}_1 \subset \bigcup_{i \in I_1} Q_i^1$, and we set $A_1 := \bigcup_{i\in I_1} Q_i^1$. We operate in an analogous way at each step k of the construction.

The estimate of $\int_{Q_0} |N^h\sigma|^p w\, d\mu$. Here there are little changes too. Equation (9.3) is proved inductively in the same way. Let us see the required modifications in the first step. The definition of B_0 is different now: $B_0 := \{x \in Q_0 \setminus \mathcal{ST}_{h+5} : S_{h+3}\sigma(x) \le \varepsilon\lambda_0\}$. With this new definition, (9.5) holds. On the other hand, notice that $\int_{Q_0 \cap \mathcal{ST}_{h+5}} |N^{h+20}\sigma|^p w\, d\mu = 0$, since $N^{h+20}\sigma(x) = 0$ if $x \in \mathcal{ST}_{h+5}$.

The definition of D_0 does not change, and all the other estimates remain valid. In particular, (9.8) holds now too, because $N^{H_x^1+20}\sigma(x) = 0$ if $x \in \mathcal{AS}_1$ (recall that the definition of A_1 has changed).

The changes required at each step k are analogous.

The estimate of $\sum_k \sigma(A_k)$. The former arguments remain valid in the new situation.

Changes in the proof of Lemma 10.1 and Lemma 10.2. The proof of Lemma 10.1 does not change. In the arguments for Lemma 10.2, we have to take into account that if $x \in B_k$ and δ is small enough, then $x \notin \mathcal{ST}_{h+M-1}$. Indeed, if $x \in Q_i$, with $Q_i \in \mathcal{AD}_h$, then we have $T(f\chi_{U_i^k})(x) \geq \frac{\alpha-1}{\alpha}\alpha^{k+2}$, and $S_j(f\chi_{U_i^k})(x) \leq \delta\alpha^k$ for $h-M \leq j \leq h+M$. These inequalities imply $S_j(f\chi_{U_i^k})(x) > \varepsilon_0\alpha^k \neq 0$ for some $j \geq M+1$ if δ is small enough. In particular, $x \notin \mathcal{ST}_{h+M-1}$.

If we assume $M \geq 20$, for instance, then all the cubes P_x that appear in the proof of Lemma 10.2 exist and are transit cubes, and the same estimates hold.

12. Relationship with $RBMO(\mu)$ and final remarks

Let us recall one of the equivalent definitions of the space $RBMO(\mu)$ introduced in [**To2**]. We say that $f \in L^1_{\text{loc}}(\mu)$ belongs to $RBMO(\mu)$ if there exists a collection of numbers $\{f_Q\}_{Q\subset\mathbb{R}^d} \subset \mathbb{R}$ such that

$$\int_Q |f(x) - f_Q|\, d\mu(x) \leq C_f\, \mu(2Q)$$

for each cube $Q \subset \mathbb{R}^d$ and

$$|f_Q - f_R| \leq C_f(1 + \delta(Q,R)) \tag{12.1}$$

for all the cubes Q, R with $Q \subset R$. The optimal constant C_f is the $RBMO(\mu)$ norm of f, which we denote by $\|f\|_*$.

Let $1 < p < \infty$. In general, if $w \in Z_p$, then $\log w \notin RBMO(\mu)$. This follows easily from Example 2.3. Indeed, in this case it can be checked that $\delta(0, I_k) \leq C$ uniformly on k. As a consequence, for all $f \in RBMO(\mu)$, the numbers f_{I_k} are bounded uniformly on k. If moreover f is constant on each interval I_k, then we deduce $f \in L^\infty(\mu)$. However, the weight w_0 of Example 2.3 is constant on each interval I_k and it is not a bounded function, and the same happens with $\log w_0$.

On the other hand, if $f \in RBMO(\mu)$, then there exists some $\varepsilon > 0$ depending on $\|f\|_*$, p such that $e^{\varepsilon f} \in Z_p$. To prove this, first we will show in the following proposition that a weight of the type $e^{\varepsilon f}$, with $f \in RBMO(\mu)$, satisfies a (rather strong) property in the spirit of the classical A_p condition.

Proposition 12.1. *Let $1 < p < \infty$. If $f \in RBMO(\mu)$ and $\varepsilon = \varepsilon(\|f\|_*, p) > 0$ is small enough, then*

$$\frac{1}{\mu(2Q)} \int_Q e^{\varepsilon f}\, d\mu \cdot \left[\frac{1}{\mu(2R)} \int_R e^{-\varepsilon f p'/p}\, d\mu\right]^{p/p'} \leq C_{32} e^{C_{33}\delta(Q,R)}, \tag{12.2}$$

for all the cubes Q, R with $Q \subset R$ or $R \subset Q$, where C_{32}, C_{33} are positive constants depending on n, d, C_0.

Proof: The funcions $f \in RBMO(\mu)$ satisfy an inequality of John-Nirenberg type (see [**To2**, Theorem 3.1]), which implies that for some constants C_{34}, C_{35} and any cube Q and $\lambda > 0$ we have

$$\int_Q \exp(C_{34}|f(x) - f_Q|/\|f\|_*)\, d\mu(x) \leq C_{35}\mu(2Q).$$

If we take $\varepsilon \leq C_{34} \min(1, p/p')/\|f\|_*)$ and we use (12.1), we deduce (12.2). □

Remark 12.2. For $1 < p < \infty$, in Lemma 4.1, the statement (e) can be replaced by the following weaker assumption:

(e') For all $k \in \mathbb{Z}$ and all cubes Q,

$$\int_Q |S_k(w\chi_Q)|^{p'} \sigma\, d\mu \leq C\, w(2Q), \tag{12.3}$$

with C independent of k and Q.

To see this, only some minor changes (which are left for the reader) in the proof of Lemma 4.1 are required.

Since (e) and (e') in Lemma 4.1 are equivalent, we deduce that the statement (e) of Lemma 4.2 can be weakened in the analogous way: We only need to compute both integrals over Q, and on the right hand side Q can be replaced by $2Q$.

Theorem 12.3. *Let $1 < p < \infty$. If $f \in RBMO(\mu)$ and $\varepsilon = \varepsilon(\|f\|_*, p) > 0$ is small enough, then $e^{\varepsilon f} \in Z_p$.*

Proof: By the preceding remark, we only need to show that $w := e^{\varepsilon f}$ satisfies (12.3) and its corresponding dual estimate. Moreover, for simplicity we will assume that there are no stopping cubes.

Let us see that (12.3) holds any given cube $Q \in \mathcal{AD}_h$. We may assume $k \geq h - 3$, since $S_k(w\chi_Q)(x) \leq C \sum_{i=-3}^{3} S_{h+i}(w\chi_Q)(x)$ for $x \in Q \cap \operatorname{supp}(\mu)$.

For each $x \in Q \cap \operatorname{supp}(\mu)$, let R_x be a doubling cube centered at x, with $R_x \in \mathcal{AD}_{k+10,k+9}$. Let $\bigcup_i R_i \supset Q \cap \operatorname{supp}(\mu)$ be a Besicovitch covering of $Q \cap \operatorname{supp}(\mu)$ with this type of cubes. Notice that $R_i \subset 2Q$ for all i. Let $Q_x \in \mathcal{AD}_{k,k-2}$ be some cube centered at x. If $x \in R_i$, then $R_i \subset Q_x$ because $\ell(R_i) \ll \ell(Q_x)$. Since $\delta(R_i, Q_x) \leq C$, by (12.2) we have

$$\frac{1}{\mu(2Q_x)} \int_{Q_x} w\, d\mu \cdot \left[\frac{1}{\mu(R_i)} \int_{R_i} \sigma\, d\mu\right]^{p/p'} \leq C.$$

Taking a suitable mean over cubes Q_x centered at x (as in the proof of Lemma 3.11), we obtain

$$S_k w(x) \cdot \left[m_{R_i}(\sigma)\right]^{p/p'} \leq C,$$

for all $x \in Q$. Then we get

$$\begin{aligned}\int_Q |S_k(w\chi_Q)|^{p'} \sigma\, d\mu &\leq \sum_i \int_{R_i} |S_k(w)|^{p'} \sigma\, d\mu \\ &\leq C \sum_i \frac{\sigma(R_i)}{(m_{R_i}\sigma)^p} = C \sum_i \frac{\mu(R_i)}{(m_{R_i}\sigma)^{p-1}}.\end{aligned}$$

By Hölder's inequality, $1 \leq m_{R_i} w \cdot (m_{R_i}\sigma)^{p-1}$. Thus,

$$\int_Q |S_k(w\chi_Q)|^{p'} \sigma\, d\mu \leq C \sum_i m_{R_i} w \cdot \mu(R_i) = C \sum_i w(R_i) \leq C\, w(2Q).$$

The estimate dual to (12.3) is proved in an analogous way. □

We will finish with some remarks and open questions:

Remark 12.4. (a) Using Lemma 9.2 and modifying a little the proof of the implication (e) ⇒ (c) of Lemma 4.1 one can show that $w \in Z_p^{\text{weak}}$ if and only if there exists some $\lambda > 1$ such that $\int N(w\chi_Q)^p\, \sigma\, d\mu \leq C\, w(\lambda Q)$ for all cubes Q. We don't know if this holds with $\lambda = 1$ too.

(b) We don't know if $Z_p = Z_p^{\text{weak}}$.

(c) In the case $p = 1$, a statement such as (e) in Lemma 4.1 is missing. We don't know if there is a reasonable substitute.

References

[CF] R. R. Coifman and C. Fefferman, Weighted norm inequalities for maximal functions and singular integrals, *Studia Math.* **51** (1974), 241–250.

[GM] J. García-Cuerva and J. M. Martell, Two-weight norm inequalities for maximal operators and fractional integrals on non-homogeneous spaces, *Indiana Univ. Math. J.* **50(3)** (2001), 1241–1280.

[Gu] M. de Guzmán, *"Differentiation of integrals in R^n"*, with appendices by Antonio Córdoba, and Robert Fefferman, and two by Roberto Moriyón, Lecture Notes in Mathematics **481**, Springer-Verlag, Berlin-New York, 1975.

[HMW] R. Hunt, B. Muckenhoupt and R. Wheeden, Weighted norm inequalities for the conjugate function and Hilbert transform, *Trans. Amer. Math. Soc.* **176** (1973), 227–251.

[Ko] Y. Komori, Weighted estimates for operators generated by maximal functions on nonhomogeneous spaces, *Georgian Math. J.* **12(1)** (2005), 121–130.

[MMNO] J. Mateu, P. Mattila, A. Nicolau and J. Orobitg, BMO for nondoubling measures, *Duke Math. J.* **102(3)** (2000), 533–565.

[Mo] A. P. Morse, Perfect blankets, *Trans. Amer. Math. Soc.* **61** (1947), 418–442.

[Mu] B. Muckenhoupt, Weighted norm inequalities for the Hardy maximal function, *Trans. Amer. Math. Soc.* **165** (1972), 207–226.

[MW] B. Muckenhoupt and R. L. Wheeden, Some weighted weak-type inequalities for the Hardy-Littlewood maximal function and the Hilbert transform, *Indiana Univ. Math. J.* **26(5)** (1977), 801–816.

[NTV1] F. Nazarov, S. Treil and A. Volberg, Cauchy integral and Calderón-Zygmund operators on nonhomogeneous spaces, *Internat. Math. Res. Notices* **15** (1997), 703–726.

[NTV2] F. Nazarov, S. Treil and A. Volberg, Weak type estimates and Cotlar inequalities for Calderón-Zygmund operators on nonhomogeneous spaces, *Internat. Math. Res. Notices* **9** (1998), 463–487.

[OP] J. Orobitg and C. Pérez, A_p weights for nondoubling measures in $\mathbb{R}^n$ and applications, *Trans. Amer. Math. Soc.* **354(5)** (2002), 2013–2033 (electronic).

[Pé] C. Pérez, Weighted norm inequalities for singular integral operators, *J. London Math. Soc. (2)* **49(2)** (1994), 296–308.

[Sak] E. Saksman, The local mapping properties of the Hilbert transform, Preprint 229, University of Helsinki, Department of Mathematics (1999).

[Saw1] E. Sawyer, A two weight weak type inequality for fractional integrals, *Trans. Amer. Math. Soc.* **281(1)** (1984), 339–345.

[Saw2] E. Sawyer, A characterization of two weight norm inequalities for fractional and Poisson integrals, *Trans. Amer. Math. Soc.* **308(2)** (1988), 533–545.

[St] E. M. Stein, *"Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals"*, With the assistance of Timothy S. Murphy, Princeton Mathematical Series **43**, Monographs in Harmonic Analysis III, Princeton University Press, Princeton, NJ, 1993.

[To1] X. Tolsa, L^2-boundedness of the Cauchy integral operator for continuous measures, *Duke Math. J.* **98(2)** (1999), 269–304.

[To2] X. Tolsa, BMO, H^1, and Calderón-Zygmund operators for non doubling measures, *Math. Ann.* **319(1)** (2001), 89–149.

[To3] X. Tolsa, The space H^1 for nondoubling measures in terms of a grand maximal operator, *Trans. Amer. Math. Soc.* **355(1)** (2003), 315–348 (electronic).

[To4] X. Tolsa, Littlewood-Paley theory and the $T(1)$ theorem with non-doubling measures, *Adv. Math.* **164(1)** (2001), 57–116.

[Vo] A. Volberg, *"Calderón-Zygmund capacities and operators on nonhomogeneous spaces"*, CBMS Regional Conference Series in Mathematics **100**, Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 2003.

Institució Catalana de Recerca i Estudis Avançats (ICREA) and
Departament de Matemàtiques
Universitat Autònoma de Barcelona
08193 Bellaterra (Barcelona)
Spain
E-mail address: xtolsa@mat.uab.cat

Primera versió rebuda el 2 d'octubre de 2006,
darrera versió rebuda el 13 de febrer de 2007.